

**Figure 1: 108P5H8 SSH sequence of 448 nucleotides (SEQ ID: \_\_\_\_\_)**

```

1 GATCCAGATT TCTCTGCACA CTGGACTTCG TAGAGTAAGT GTGGTAGACA AAGAGACTAC
61 ACTGCACAAC CACCAGTGAA TATCATTGCT AAGAAGACTT TGGGTCGTGT TTCTCAGCCA
121 CTCTCACAGC TTTTGTAGAC TTATTTGATT TTGAAACAAG CAGTTAGCTA AATCTATTTT
181 CCTTTTATGC ATATATGTTA ATTGGCTCAA CTTAATATGG TGTTCCTTACA GAATATGAGC
241 CCATTTGAAA TAAGGTTTTA GGCAATTTTG CTGTTGGCTC TGATTTGTAT ATAGCAAATT
301 TAAAGGTACA GAGTGTTTTCC TAGATAGAAG ATTAGTTCAT TTGGTTCATT TTGTCTTTGA
361 AGCAAGCCAA GCTCATGAGC CAGTTGGTTA TTTGTCATAA ATGAACACCC ATCACTATAT
421 GCTATGTTGA GGGGAGGCAA GGCTGATC

```

**Figure 2A. The cDNA (SEQ ID. NO. :\_\_\_\_) and amino acid sequence (SEQ ID. NO. :\_\_\_\_) of 108P5H8 v.1.** The start methionine is underlined. The open reading frame extends from nucleic acid 253-1542 including the stop codon.

```

1 gccggcctccagcagcgggcgcgggcgcgagcagcagcccccactctcctgcgggccgcg
61 ggtggagcagcgcgagccccgcctcgctgagccggccggggcggggagatgagttgcggc
121 cccgcggcagcgcggccaggatggggagggagcgcgcgccactgcctcgagaactggcgct
181 ccggtgaagtaggcgcgcggcgccgtccgcctcccccagccgttccgcaccgcggccgc
1      M A G S G A W K R L K S M L R K
241 tcagcctctgccATGGCCGGCTCTGGCGCGTGGAAGCGCCTCAAATCTATGCTAAGGAAG
17 D D A P L F L N D T S A F D F S D E A G
301 GATGATGCGCCGCTGTTTTTAAATGACACCAGCGCCTTTGACTTCTCGGATGAGGCGGGG
37 D E G L S R F N K L R V V V A D D G S E
361 GACGAGGGGCTTTCTCGGTTCAACAAACTTCGAGTTGTGGTGGCCGATGACGGTTCCGAA
57 A P E R P V N G A H P T L Q A D D D S L
421 GCCCCGAAAGGCCTGTTAACGGGGCGCACCCGACCCTCCAGGCCGACGATGATTCCTTA
77 L D Q D L P L T N S Q L S L K V D S C D
481 CTGGACCAAGACTTACCTTTGACCAACAGTCAGCTGAGTTTGAAGGTGGACTCCTGTGAC
97 N C S K Q R E I L K Q R K V K A R L T I
541 AACTGCAGCAAACAGAGAGAGATACTGAAGCAGAGAAAGGTGAAAGCCAGGTTGACCATT
117 A A V L Y L L F M I G E L V G G Y I A N
601 GCTGCCGTTCTGTACTTGCTTTTCATGATTGGAGAAGTTGTAGGTGGATACATTGCAAAT
137 S L A I M T D A L H M L T D L S A I I L
661 AGCCTAGCAATCATGACAGATGCACTTCATATGTAACTGACCTAAGCGCCATCATACTC
157 T L L A L W L S S K S P T K R F T F G F
721 ACCCTGCTTGCTTTGTGGCTATCATCAAAATCACCAACCAAAAGATTACCTTTGGATTT
177 H R L E V L S A M I S V L L V Y I L M G

```

781 CATCGCTTAGAGGTTTTGTCAGCTATGATTAGTGTGCTGTTGGTGTATATACTTATGGGA  
 197 F L L Y E A V Q R T I H M N Y E I N G D  
 841 TTCCTCTTATATGAAGCTGTGCAAAGAACTATCCATATGAACTATGAAATAAATGGAGAT  
 217 I M L I T A A V G V A V N V I M G F L L  
 901 ATAATGCTCATCACCGCAGCTGTTGGAGTTGCAGTTAATGTAATAATGGGGTTTCTGTTG  
 237 N Q S G H R H S H S H S L P S N S P T R  
 961 AACCAGTCTGGTCACCGTCACTCCCATTCCCCTCCCTGCCTTCAAATCCCCCTACCAGA  
 257 G S G C E R N H G Q D S L A V R A A F V  
 1021 GGTTCCTGGGTGTGAACGTAACCATGGGCAGGATAGCCTGGCAGTGAGAGCTGCATTTGTA  
 277 H A L G D L V Q S V G V L I A A Y I I R  
 1081 CATGCTTTGGGAGATTTGGTACAGAGTGTGGTGTGCTAATAGCTGCATACATCATACGA  
 297 F K P E Y K I A D P I C T Y V F S L L V  
 1141 TTCAAGCCAGAATACAAGATTGCTGACCCCATCTGTACATACGTATTTTCATTACTTGTG  
 317 A F T T F R I I W D T V V I I L E G V P  
 1201 GCTTTTACAACATTTTGAATCATATGGGATACAGTAGTTATAATACTAGAAGGTGTGCCA  
 337 S H L N V D Y I K E A L M K I E D V Y S  
 1261 AGCCATTTGAATGTAGACTATATCAAAGAAGCCTTGATGAAAATAGAAGATGTATATTCA  
 357 V E D L N I W S L T S G K S T A I V H I  
 1321 GTCGAAGATTTAAATATCTGGTCTCTCACTTCAGGAAAATCTACTGCCATAGTTTCACATA  
 377 Q L I P G S S S K W E E V Q S K A N H L  
 1381 CAGCTAATTCCTGGAAGTTCATCTAAATGGGAGGAAGTACAGTCCAAAGCAAACCATTTA  
 397 L L N T F G M Y R C T I Q L Q S Y R Q E  
 1441 TTATTGAACACATTTGGCATGTATAGATGTACTATTTCAGCTTCAGAGTTACAGGCAAGAA  
 417 V D R T C A N C Q S S S P \*  
 1501 GTGGACAGAACTTGTGCAAATTGTCAGAGTTCTAGTCCCTAAAttttatgtattttgggaa  
 1561 ctctctgccttatttatcctgcagtcacagacttgagagcaataaatgcaaacctaaatga  
 1621 gaaaatggaatccctgacagctgtgtccgtatcaagcatcagtcctctcaaacagttgccc  
 1681 cagcctgacagtgctagtctctgtttaatggtaaaaggagactttgccataattttcaga  
 1741 tgaagatgtttcccaaactgtttacagaatgagatgtgactctacagatacctcatag  
 1801 aagacaatccaagatcatacttcattaacttgacagagtacgtgtcttaaaggaagcatc  
 1861 aagaattcaatattttgcatttaaaaaatactttttaaggccattttatattaagccagtg  
 1921 tggaaaactgaattttttttattatgtataataatctcgacacccagcttctggaattgc  
 1981 tgctttctttttacagaaattactaccaacagatttcaggaagtactagtagttatccc  
 2041 aaaagtggaaataagcatgtattcctaagtgtttcagaaatgttttatttcacacataagt  
 2101 cttaatgtttattgtttatgattatactttataaacaaccttttcagatgctacagggttt  
 2161 tgaatctcaaagttaacattttttcattattttgtaatcttagaaccaaatctttattttatt  
 2221 gtggctcactgttattaaatgatttaggaaatactttcaatattattctgaaatggctgaag

[illegible]

**Figure 2B. The cDNA (SEQ ID. NO. :\_\_\_\_) and amino acid sequence (SEQ ID. NO. :\_\_\_\_) of 108P5H8 v.2.** The start methionine is underlined. The open reading frame extends from nucleic acid 1-1290 including the stop codon.

```

1  M A G S G A W K R L K S M L R K D D A P
1  ATGGCCCGCTCTGGCGCGTGGAAGCGCCTCAAATCTATGCTAAGGAAGGATGATGCGCCC
21 L F L N D T S A F D F S D E A G D E G L
61 CTGTTTTTAAATGACACCAGCGCCTTTGACTTCTCGGATGAGGCGGGGACGAGGGGCTT
41 S R F N K L R V V V A D D G S E A P E R
121 TCTCGGTTCAACAACTTCGAGTTGTGGTGGCCGATGACGGTTCCGAAGCCCCGAAAGG
61 P V N G A H P T L Q A D D D S L L D Q D
181 CCTGTTAACGGGGCGCACCCGACCCTCCAGGCCGACGATGATTCCTTACTGGACCAAGAC
81 L P L T N S Q L S L K V D S C D N C S K
241 TTACCTTTGACCAACAGTCAGCTGAGTTTGAAGGTGGACTCCTGTGACAACTGCAGCAAA
101 Q R E I L K Q R K V K A R L T I A A V L
301 CAGAGAGAGATACTGAAGCAGAGAAAGGTGAAAGCCAGGTTGACCATTGCTGCCGTTCTG
121 Y L L F M I G E L V G G Y I A N S L A I
361 TACTTGCTTTTCATGATTGGAGAACTTGTAGGTGGATACATTGCAAATAGCCTAGCAATC
141 M T D A L H M L T D L S A I I L T L L A
421 ATGACAGATGCACTTCATATGTAACTGACCTAAGCGCCATCATACTCACCTGCTTGCT
161 L W L S S K S P T K R F T F G F H R L E
481 TTGTGGCTATCATCAAAATCACCAACCAAAAAGATTCACCTTTGGATTTTCATCGCTTAGAG
181 V L S A M I S V L L V Y I L M G F L L Y
541 GTTTTGTGAGCTATGATTAGTGTGCTGTTGGTGTATATACTTATGGGATTCTCTTATAT
201 E A V Q R T I H M N Y E I N G D I M L I
601 GAAGCTGTGCAAAGAACTATCCATATGAACTATGAAATAAATGGAGATATAATGCTCATC
221 T A A V G V A V N V I M G F L L N Q S G
661 ACCGCAGCTGTTGGAGTTGCAGTTAATGTAATAATGGGGTTTCTGTTGAACCAGTCTGGT
241 H R H S H S H S L P S N S P T R G S G C
721 CACCGTCACTCCCATTCCTCCCTGCCTTCAAATTCCTTACCAGAGGTTCTGGGTGT
261 E R N H G Q D S L A V R A A F V H A L G
781 GAACGTAACCATGGGCAGGATAGCCTGGCAGTGAGAGCTGCATTTGTACATGCTTTGGGA
281 D L V Q S V G V L I A A Y I I R F K P E
841 GATTTGGTACAGAGTGTGGTGTGCTAATAGCTGCATACATCATACGATTCAAGCCAGAA
301 Y K I A D P I C T Y V F S L L V A F T T
901 TACAAGATTGCTGATCCCATCTGTACATACGTATTTTCATTACTTGTGGCTTTTACAACA
321 F R I I W D T V V I I L E G V P S H L N

```

961 TTTCGAATCATATGGGATACAGTAGTTATAATACTAGAAGGTGTGCCAAGCCATTTGAAT  
341 V D Y I K E A L M K I E D V Y S V E D L  
1021 GTAGACTATATCAAAGAAGCCTTGATGAAAATAGAAGATGTATATTCAGTCGAAGATTTA  
361 N I W S L T S G K S T A I V H I Q L I P  
1081 AATATCTGGTCTCTCACTTCAGGAAAATCTACTGCCATAGTTCACATACAGCTAATTCCT  
381 G S S S K W E E V Q S K A N H L L L N T  
1141 GGAAGTTCATCTAAATGGGAGGAAGTACAGTCCAAAGCAAACCATTTATTATTGAACACA  
401 F G M Y R C T I Q L Q S Y R Q E V D R T  
1201 TTTGGCATGTATAGATGTACTATTCAGCTTCAGAGTTACAGGCAAGAAGTGGACAGAACT  
421 C A N C Q S S S P \*  
1261 TGTGCAAATTGTCAGAGTTCTAGTCCCTAAAttttatgtattttggggactcctgccttat  
1321 ttatcctgcagtcacagacttgagagcaataaatgcaaacctaaatgagaaaatggaatc  
1381 cctgacagctgtgtccgtatcaagcatcagtcctctcaaacagttgccccagcctgacagt  
1441 gctagtcctctgtttaatggtaaaaggagactttgccataattttcagatgaagatgtttc  
1501 ccaaactgtttacagaatgagatgtgactctacagatacctcatag

**Figure 2C. The cDNA (SEQ ID. NO. :\_\_\_\_) and amino acid sequence (SEQ ID. NO. :\_\_\_\_) of 108P5H8 v.3.** The start methionine is underlined. The open reading frame extends from nucleic acid 1-1290 including the stop codon.

```

1 M A G S G A W K R L K S M L R K D D A P
1 ATGGCCCGGCTCTGGCGCGTGGAAGCGCCTCAAATCTATGCTAAGGAAGGATGATGCGCCG
21 L F L N D T S A F E F S D E A G D E G L
61 CTGTTTTTTAAATGACACCAGCGCCTTTGAGTTCTCGGATGAGGCGGGGACGAGGGGCTT
41 S R F N K L R V V V A D D G S E A P E R
121 TCTCGGTTCAACAAACTTCGAGTTGTGGTGGCCGATGACGGTTCCGAAGCCCCGAAAGG
61 P V N G A H P T L Q A D D D S L L D Q D
181 CCTGTTAACGGGGCGCACCCGACCCTCCAGGCCGACGATGATTCTTACTGGACCAAGAC
81 L P L T N S Q L S L K V D S C D N C S K
241 TTACCTTTGACCAACAGTCAGCTGAGTTTGAAGGTGGACTCCTGTGACAACTGCAGCAAA
101 Q R E I L K Q R K V K A R L T I A A V L
301 CAGAGAGAGATACTGAAGCAGAGAAAGGTGAAAGCCAGGTTGACCATTGCTGCCGTTCTG
121 Y L L F M I G E L V G G Y I A N S L A I
361 TACTTGCTTTTCATGATTGGAGAACTTGTAGGTGGATACATTGCAAATAGCCTAGCAATC
141 M T D A L H M L T D L S A I I L T L L A
421 ATGACAGATGCACTTCATATGTTAACTGACCTAAGCGCCATCATACTCACCTGCTTGCT
161 L W L S S K S P T K R F T F G F H R L E
481 TTGTGGCTATCATCAAAATCACCAACCAAAAGATTACCTTTGGATTTTCATCGCTTAGAG
181 V L S A M I S V L L V Y I L M G F L L Y
541 GTTTTGTGAGCTATGATTAGTGTGCTGTTGGTGTATATACTTATGGGATTCTCTTATAT
201 E A V Q R T I H M N Y E I N G D I M L I
601 GAAGCTGTGCAAAGAACTATCCATATGAACTATGAAATAAATGGAGATATAATGCTCATC
221 T A A V G V A V N V I M G F L L N Q S G
661 ACCGCAGCTGTTGGAGTTGCAGTTAATGTAATAATGGGGTTTCTGTTGAACCAAGTCTGGT
241 H R H S H S H S L P S N S P T R G S G C
721 CACCGTCACTCCCATTCCTCCCTCCCTGCTTCAAATTCCTCCCTACCAGAGGTTCTGGGTGT
261 E R N H G Q D S L A V R A A F V H A L G
781 GAACGTAACCATGGGCAGGATAGCCTGGCAGTGAGAGCTGCATTTGTACATGCTTTGGGA
281 D L V Q S V G V L I A A Y I I R F K P E
841 GATCTGGTACAGAGTGTGGTGTGCTAATAGCTGCATACATCATAACGATTCAAGCCAGAA
301 Y K I A D P I C T Y V F S L L V A F T T
901 TACAAGATTGCTGACCCCATCTGTACATACGTATTTTCATTACTTGTGGCTTTTACAACA

```

321 F R I I W D T V V I I L E G V P S H L N  
 961 TTTCGAATCATATGGGATACAGTAGTTATAATACTAGAAGGTGTGCCAAGCCATTTGAAT  
 341 V D Y I K E A L M K I E D V Y S V E D L  
 1021 GTAGACTATATCAAAGAAGCCTTGATGAAAATAGAAGATGTATATTCAGTCGAAGATTTA  
 361 N I W S L T S G K S T A I V H I Q L I P  
 1081 AATATCTGGTCTCTCACTTCAGGAAAATCTACTGCCATAGTTCACATACAGCTAATTCCT  
 381 G S S S K W E E V Q S K A N H L L L N T  
 1141 GGAAGTTCATCTAAATGGGAGGAAGTACAGTCCAAAGCAAACCATTTATTATTGAACACA  
 401 F G M Y R C T I Q L Q S Y R Q E V D R T  
 1201 TTTGGCATGTATAGATGTACTATTTCAGCTTCAGAGTTACAGGCAAGAAGTGGACAGAACT  
 421 C A N C Q S S S P \*  
 1261 TGTGCAAATTGTCAGAGTTCTAGTCCCTAAAttttatgtattgttttagcattgctgaatt  
 1321 cactttatttatcctgcagtcacagacttgagagcaataaatgcaaacctaaatgagaaa  
 1381 atggaatccctgacagctgtgtccgtatcaagcatcagtcctctcaaacagttgccccagc  
 1441 ctgacagtgctagtctctgtttaatggtaaaaggagactttgccataattttcagatgaa  
 1501 gatgtttcccaaactgttttacagaatgagatgtgactcctacagatacctcatag

**Figure 3:**

**Figure 3A. Amino acid sequence corresponding to 108P5H8 v.1 and 108P5H8 v.2 (SEQ ID. NO. : \_\_\_\_).** The 108P5H8 v.1 and 108P5H8 v.2 proteins have 429 amino acids.

```

1  MAGSGAWKRL  KSMRLRKDDAP  LFLNDTSAFD  FSDEAGDEGL  SRFNKL RVVV  ADDGSEAPER
61  PVNGAHPTLQ  ADDDSLDDQD  LPLTNSQLSL  KVDSCDNCSK  QREILKQRKV  KARLTIAAVL
121 YLLFMIGELV  GGYIANSLAI  MTDALHMLTD  LSAILLTLLA  LWLSSKSPTK  RFTFGFHRLE
181 VLSAMISVLL  VYILMGFLLY  EAVQRTIHMN  YEINGDIMLI  TAAVGVA VNV  IMGFLLNQSG
241 HRHSHSHSLP  SNSPTRGSGC  ERNHGQDSL A  VRAAFVHALG  DLVQSVGVLI  AAYIIRFKPE
301 YKIADPICTY  VFSLLVAF TT  FRIIWDTVVI  ILEGVPSHLN  VDYIKEALMK  IEDVYSVEDL
361 NIWSLTSGKS  TAIVHIQLIP  GSSSKWEEVQ  SKANHLLLNT  FGMYRCTIQL  QSYRQEVDR T
421 CANCQSSSP*

```

**Figure 3B. Amino acid sequence corresponding to 108P5H8 v.3 (SEQ ID. NO. : \_\_\_\_).**

The 108P5H8 v.3 protein has 429 amino acids.

```

1  MAGSGAWKRL  KSMRLRKDDAP  LFLNDTSAFE  FSDEAGDEGL  SRFNKL RVVV  ADDGSEAPER
61  PVNGAHPTLQ  ADDDSLDDQD  LPLTNSQLSL  KVDSCDNCSK  QREILKQRKV  KARLTIAAVL
121 YLLFMIGELV  GGYIANSLAI  MTDALHMLTD  LSAILLTLLA  LWLSSKSPTK  RFTFGFHRLE
181 VLSAMISVLL  VYILMGFLLY  EAVQRTIHMN  YEINGDIMLI  TAAVGVA VNV  IMGFLLNQSG
241 HRHSHSHSLP  SNSPTRGSGC  ERNHGQDSL A  VRAAFVHALG  DLVQSVGVLI  AAYIIRFKPE
301 YKIADPICTY  VFSLLVAF TT  FRIIWDTVVI  ILEGVPSHLN  VDYIKEALMK  IEDVYSVEDL
361 NIWSLTSGKS  TAIVHIQLIP  GSSSKWEEVQ  SKANHLLLNT  FGMYRCTIQL  QSYRQEVDR T
421 CANCQSSSP*

```



**Figure 4:**

**Figure 4A** Nucleic acid sequence alignment of the 3 variants of 108P5H8. Highlighted in yellow and underlined are the variations between the variants. The ORF extends from nucleotides 253-1542 of 108P5H8 v.1.

```

1      15 16      30 31      45 46      60 61      75 76
v.1 GCCGGCCTCCAGCAG CGGGCGGGCGGGGCG CGAGCACGACCCAC TCTCCTGCGGCCGCG GGTGGAGCAGCGCGA
GCCCGCCTCGCTGAG
v.2 -----
-----
v.3 -----
-----
91     105 106     120 121     135 136     150 151     165 166
v.1 CCGGCCGGGGCGGG GAGATGAGTTGCGGC CCCGCGGCAGCGCCC CAGGATGGGGAGGGA CGCGCGGCACTGCCC
TCGAGAACTGGCGCT
v.2 -----
-----
v.3 -----
-----
181    195 196    210 211    225 226    240 241    255 256
v.1 CCGGTGAAGTAGGCG CCGCCGGCCGTCGCG CTCCCCAAGCCGTT CCGCACCGGGCCGC TCAGCCTCTGCCATG
GCCGGCTCTGGCGCG
v.2 -----
-----
v.3 -----
-----
270    285 286    300 301    315 316    330 331    345 346
v.1 CCGGTGAAGTAGGCG CCGCCGGCCGTCGCG CTCCCCAAGCCGTT CCGCACCGGGCCGC TCAGCCTCTGCCATG
GCCGGCTCTGGCGCG
v.2 -----
-----
v.3 -----
-----
```



v.2 AACTGCAGCAACAG AGAGAGATACTGAAG CAGAGAAAGGTGAAA GCCAGGTTGACCAAT GCTGCCGTTCTGTAC  
TTGCTTTTCATGATT  
v.3 AACTGCAGCAACAG AGAGAGATACTGAAG CAGAGAAAGGTGAAA GCCAGGTTGACCAAT GCTGCCGTTCTGTAC  
TTGCTTTTCATGATT

631 645 646 660 661 675 676 690 691 705 706

720

v.1 GGAGAACTTGTAGGT GGATACATTGCAAAAT AGCCTAGCAATCATG ACAGATGCACCTTCAT ATGTTAACTGACCTA  
AGCGCCATCATACTC  
v.2 GGAGAACTTGTAGGT GGATACATTGCAAAAT AGCCTAGCAATCATG ACAGATGCACCTTCAT ATGTTAACTGACCTA  
AGCGCCATCATACTC  
v.3 GGAGAACTTGTAGGT GGATACATTGCAAAAT AGCCTAGCAATCATG ACAGATGCACCTTCAT ATGTTAACTGACCTA  
AGCGCCATCATACTC

721 735 736 750 751 765 766 780 781 795 796

810

v.1 ACCCTGCTTGCTTTG TGGCTATCATCAAAA TCACCAACC AAAAGA TTCACCTTTGGATT CATCGCTTAGAGGTT  
TTGTCAGCTATGATT  
v.2 ACCCTGCTTGCTTTG TGGCTATCATCAAAA TCACCAACC AAAAGA TTCACCTTTGGATT CATCGCTTAGAGGTT  
TTGTCAGCTATGATT  
v.3 ACCCTGCTTGCTTTG TGGCTATCATCAAAA TCACCAACC AAAAGA TTCACCTTTGGATT CATCGCTTAGAGGTT  
TTGTCAGCTATGATT

811 825 826 840 841 855 856 870 871 885 886

900

v.1 AGTGTGCTGTTGGTG TATATACTTATGGGA TTCTCTTTATATGAA GCTGTGCAAAAGAACT ATCCATATGAACTAT  
GAAATAAATGGAGAT  
v.2 AGTGTGCTGTTGGTG TATATACTTATGGGA TTCTCTTTATATGAA GCTGTGCAAAAGAACT ATCCATATGAACTAT  
GAAATAAATGGAGAT  
v.3 AGTGTGCTGTTGGTG TATATACTTATGGGA TTCTCTTTATATGAA GCTGTGCAAAAGAACT ATCCATATGAACTAT  
GAAATAAATGGAGAT

[illegible]

v.2 ATCTGTACATACGTA TTTTCATTACTTGTG GCTTTTACAACATTT CGAATCATATGGGAT ACAGTAGTTATAATA  
 CTAGAAGGTGTGCCA  
 v.3 ATCTGTACATACGTA TTTTCATTACTTGTG GCTTTTACAACATTT CGAATCATATGGGAT ACAGTAGTTATAATA  
 CTAGAAGGTGTGCCA

1261 1275 1276 1290 1291 1305 1306 1320 1321 1335  
 1336 1350  
 v.1 AGCCATTTGAATGTA GACTATATCAAAGAA GCCTTGATGAAAATA GAAGATGTATATTCA GTCGAAGATTTAAAT  
 ATCTGGTCTCTCACT  
 v.2 AGCCATTTGAATGTA GACTATATCAAAGAA GCCTTGATGAAAATA GAAGATGTATATTCA GTCGAAGATTTAAAT  
 ATCTGGTCTCTCACT  
 v.3 AGCCATTTGAATGTA GACTATATCAAAGAA GCCTTGATGAAAATA GAAGATGTATATTCA GTCGAAGATTTAAAT  
 ATCTGGTCTCTCACT

1351 1365 1366 1380 1381 1395 1396 1410 1411 1425  
 1426 1440  
 v.1 TCAGGAAAATCTACT GCCATAGTTTCACATA CAGCTAAATTCCTGGA AGTTTCATCTAAATGG GAGGAAGTACAGTCC  
 AAAGCAAACCATTTA  
 v.2 TCAGGAAAATCTACT GCCATAGTTTCACATA CAGCTAAATTCCTGGA AGTTTCATCTAAATGG GAGGAAGTACAGTCC  
 AAAGCAAACCATTTA  
 v.3 TCAGGAAAATCTACT GCCATAGTTTCACATA CAGCTAAATTCCTGGA AGTTTCATCTAAATGG GAGGAAGTACAGTCC  
 AAAGCAAACCATTTA

1441 1455 1456 1470 1471 1485 1486 1500 1501 1515  
 1516 1530  
 v.1 TTATTGAACACATTT GGCATGTATAGATGT ACTATTAGCTTCAG AGTTACAGGCAAGAA GTGGACAGAACTTGT  
 GCAAAATTGTCAGAGT  
 v.2 TTATTGAACACATTT GGCATGTATAGATGT ACTATTAGCTTCAG AGTTACAGGCAAGAA GTGGACAGAACTTGT  
 GCAAAATTGTCAGAGT  
 v.3 TTATTGAACACATTT GGCATGTATAGATGT ACTATTAGCTTCAG AGTTACAGGCAAGAA GTGGACAGAACTTGT  
 GCAAAATTGTCAGAGT

1606 1531 1545 1546 1560 1561 1575 1576 1590 1591 1605  
 1620  
 v.1 TCTAGTCCCCTAAATTT TATGTATT--TTGGG AACTCCTG-----C CTTATTATTATCCTGCA GTCACAGACTTGAGA  
 GCAATAAATGCAAAC  
 v.2 TCTAGTCCCCTAAATTT TATGTATT--TTGGG GACTCCTG-----C CTTATTATTATCCTGCA GTCACAGACTTGAGA  
 GCAATAAATGCAAAC  
 v.3 TCTAGTCCCCTAAATTT TATGTATTGTTTATG CATTGCTGAATTCA TTTATTATTATCCTGCA GTCACAGACTTGAGA  
 GCAATAAATGCAAAC

1696 1621 1635 1636 1650 1651 1665 1666 1680 1681 1695  
 1710  
 v.1 CTAAATGAGAAAAATG GAATCCCTGACAGCT GTGTCCGTATCAAGC ATCAGTCTCTCAAAC AGTTGCCCCCAGCCTG  
 ACAGTGCTAGTCTCT  
 v.2 CTAAATGAGAAAAATG GAATCCCTGACAGCT GTGTCCGTATCAAGC ATCAGTCTCTCAAAC AGTTGCCCCCAGCCTG  
 ACAGTGCTAGTCTCT  
 v.3 CTAAATGAGAAAAATG GAATCCCTGACAGCT GTGTCCGTATCAAGC ATCAGTCTCTCAAAC AGTTGCCCCCAGCCTG  
 ACAGTGCTAGTCTCT

1786 1711 1725 1726 1740 1741 1755 1756 1770 1771 1785  
 1800  
 v.1 GTTTAATGGTAAAAG GAGACTTTGCCATAA TTTTCAGATGAAAGAT GTTTCCCAAACACTG TTTACAGAAATGAGAT  
 GTGACTC-TACAGAT  
 v.2 GTTTAATGGTAAAAG GAGACTTTGCCATAA TTTTCAGATGAAAGAT GTTTCCCAAACACTG TTTACAGAAATGAGAT  
 GTGACTC-TACAGAT  
 v.3 GTTTAATGGTAAAAG GAGACTTTGCCATAA TTTTCAGATGAAAGAT GTTTCCCAAACACTG TTTACAGAAATGAGAT  
 GTGACTCCTACAGAT

1876 1801 1815 1816 1830 1831 1845 1846 1860 1861 1875  
 1890  
 v.1 ACCTCATAGAAGACA ATCCAAGATCATACT TCATTAACTTGACAG AGTACGTGTCTTAAA GGAAGCATCAAGAAT  
 TCAATATTGCATTT



[illegible]



Figure 4B Amino Acid Alignment of the 3 variants of 108P5H8

76	1	15	16	30	31	45	46	60	61	75
	v.1	MAGSGAWKRLKSMRLR	KDDAPLFLNDTSAFD	FSDEAGDEGLSRFNK	LRVVVADDGSEAPER	PVNGAHP	TLQADDD			
		LLDQDLPLTNSQLSL								
	v.2	MAGSGAWKRLKSMRLR	KDDAPLFLNDTSAFD	FSDEAGDEGLSRFNK	LRVVVADDGSEAPER	PVNGAHP	TLQADDD			
		LLDQDLPLTNSQLSL								
	v.3	MAGSGAWKRLKSMRLR	KDDAPLFLNDTSAFE	FSDEAGDEGLSRFNK	LRVVVADDGSEAPER	PVNGAHP	TLQADDD			
		LLDQDLPLTNSQLSL								
	91	105	106	120	121	135	136	150	151	165
	v.1	KVDSCDNCCKQREIL	KQKVKARLTIAAVL	YLLFMIGELVGGYIA	NSLAIMTDALHMLTD	LSAIILTLLALWLSS				
		KSPTKRFTFGFHRLE								
	v.2	KVDSCDNCCKQREIL	KQKVKARLTIAAVL	YLLFMIGELVGGYIA	NSLAIMTDALHMLTD	LSAIILTLLALWLSS				
		KSPTKRFTFGFHRLE								
	v.3	KVDSCDNCCKQREIL	KQKVKARLTIAAVL	YLLFMIGELVGGYIA	NSLAIMTDALHMLTD	LSAIILTLLALWLSS				
		KSPTKRFTFGFHRLE								
	181	195	196	210	211	225	226	240	241	255
	v.1	VLSAMISVLLVYILM	GFLLYEAVQRTIHMN	YEINGDIMLITAAVG	VAVNVIMGFLLNQSG	HRHSHSHSLPSNSPT				
		RGSGCERNHGQDSL								
	v.2	VLSAMISVLLVYILM	GFLLYEAVQRTIHMN	YEINGDIMLITAAVG	VAVNVIMGFLLNQSG	HRHSHSHSLPSNSPT				
		RGSGCERNHGQDSL								
	v.3	VLSAMISVLLVYILM	GFLLYEAVQRTIHMN	YEINGDIMLITAAVG	VAVNVIMGFLLNQSG	HRHSHSHSLPSNSPT				
		RGSGCERNHGQDSL								
	271	285	286	300	301	315	316	330	331	345
	346									

v.1 VRAAFVHALGDLVQS VGVLIAAYIIRFKPE YKIADP ICTYVFSLL VAF TTFRIIWDTVVI ILEGVP SHLNVDYIK  
EALMKIEDVYSVEDL  
v.2 VRAAFVHALGDLVQS VGVLIAAYIIRFKPE YKIADP ICTYVFSLL VAF TTFRIIWDTVVI ILEGVP SHLNVDYIK  
EALMKIEDVYSVEDL  
v.3 VRAAFVHALGDLVQS VGVLIAAYIIRFKPE YKIADP ICTYVFSLL VAF TTFRIIWDTVVI ILEGVP SHLNVDYIK  
EALMKIEDVYSVEDL

361	375	376	390	391	405	406	420	421	
v.1	NIWSLTSGKSTAIVH	IQLIPGSSSKWEEVQ	SKANHLLLLNTFGMYR	CTIQLQSYRQEVDRT	CANCQSSSP				429
v.2	NIWSLTSGKSTAIVH	IQLIPGSSSKWEEVQ	SKANHLLLLNTFGMYR	CTIQLQSYRQEVDRT	CANCQSSSP				429
v.3	NIWSLTSGKSTAIVH	IQLIPGSSSKWEEVQ	SKANHLLLLNTFGMYR	CTIQLQSYRQEVDRT	CANCQSSSP				429

Figure 5: 108P5H8 Hydrophilicity profile  
 (Hopp T.P., Woods K.R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

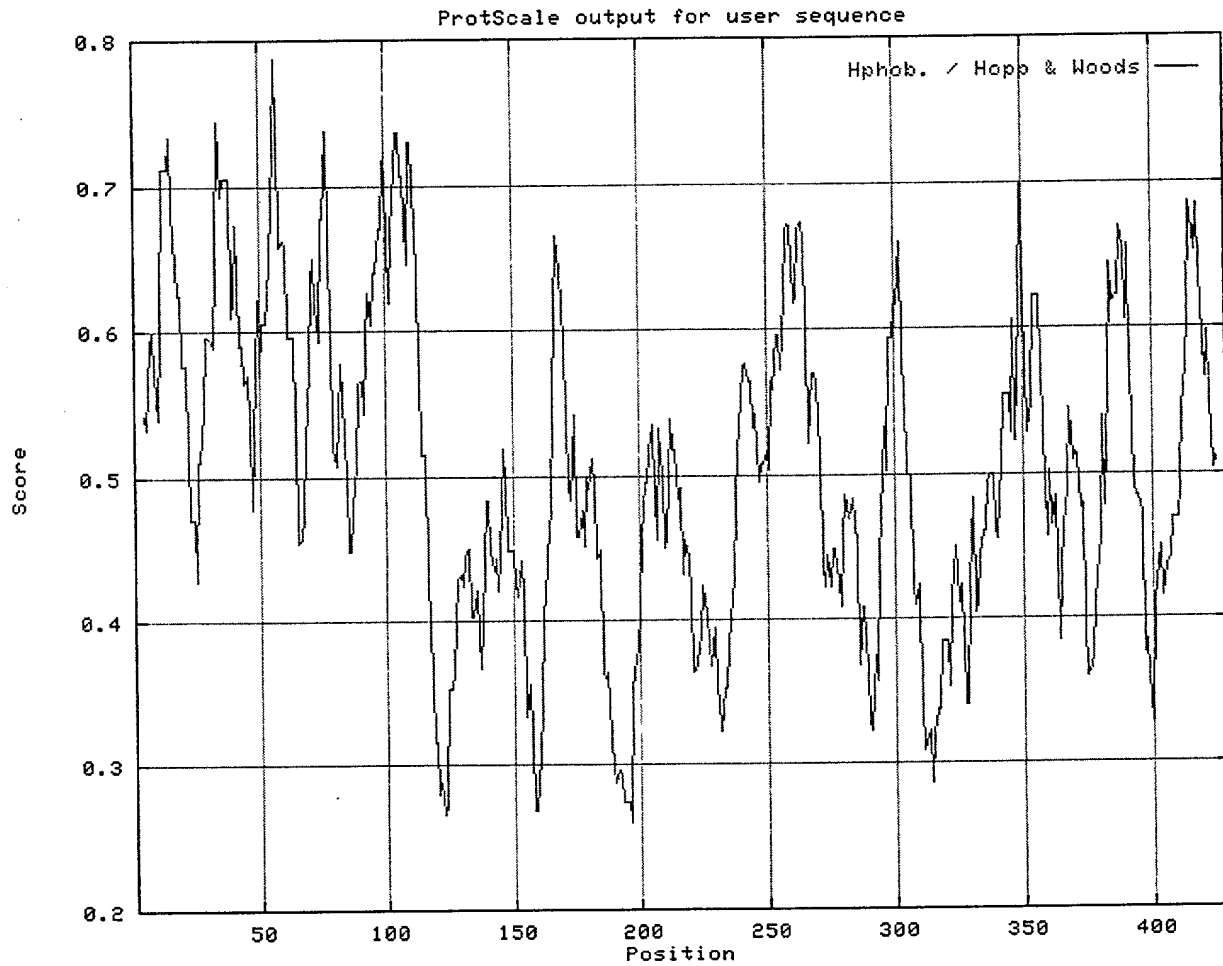
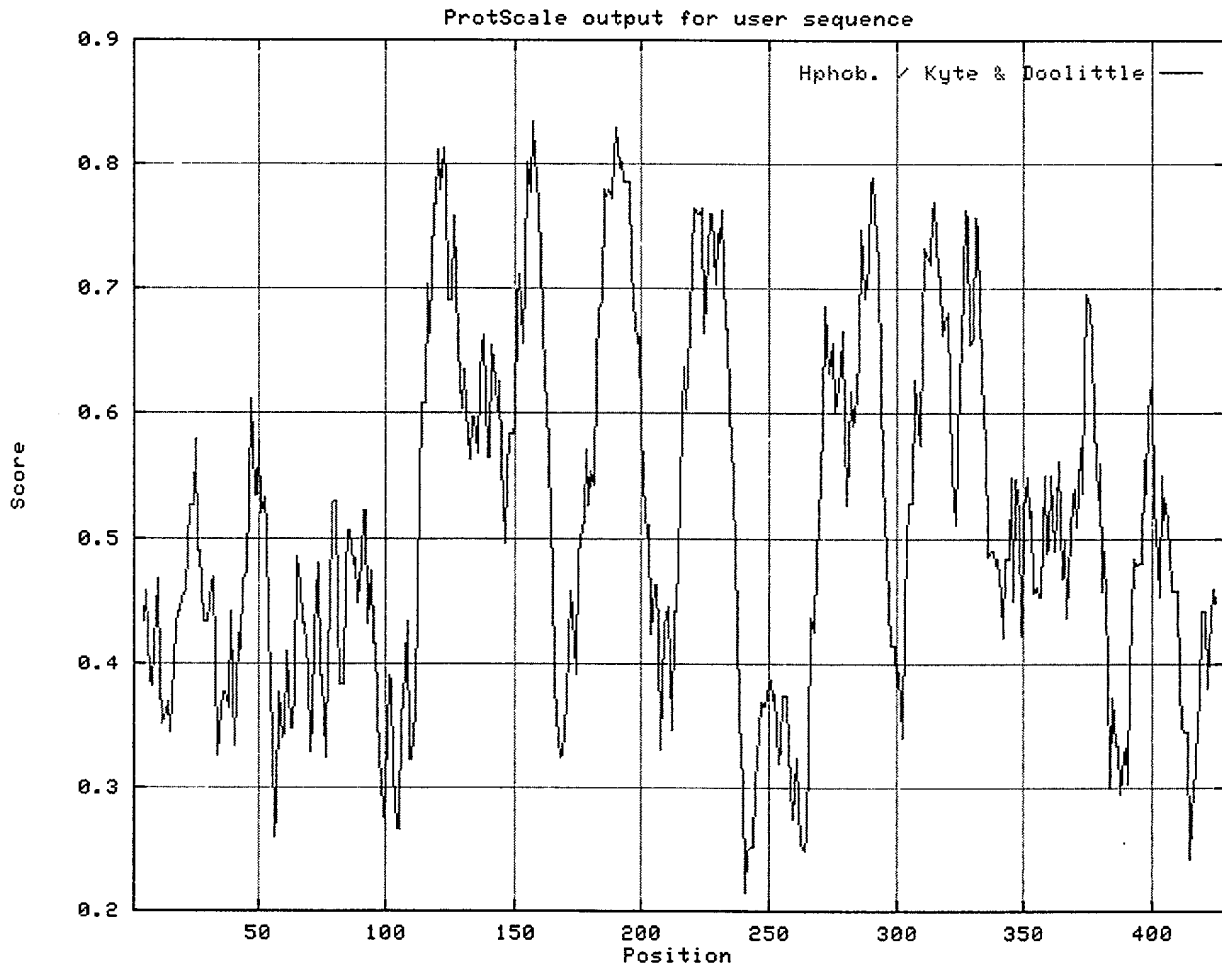
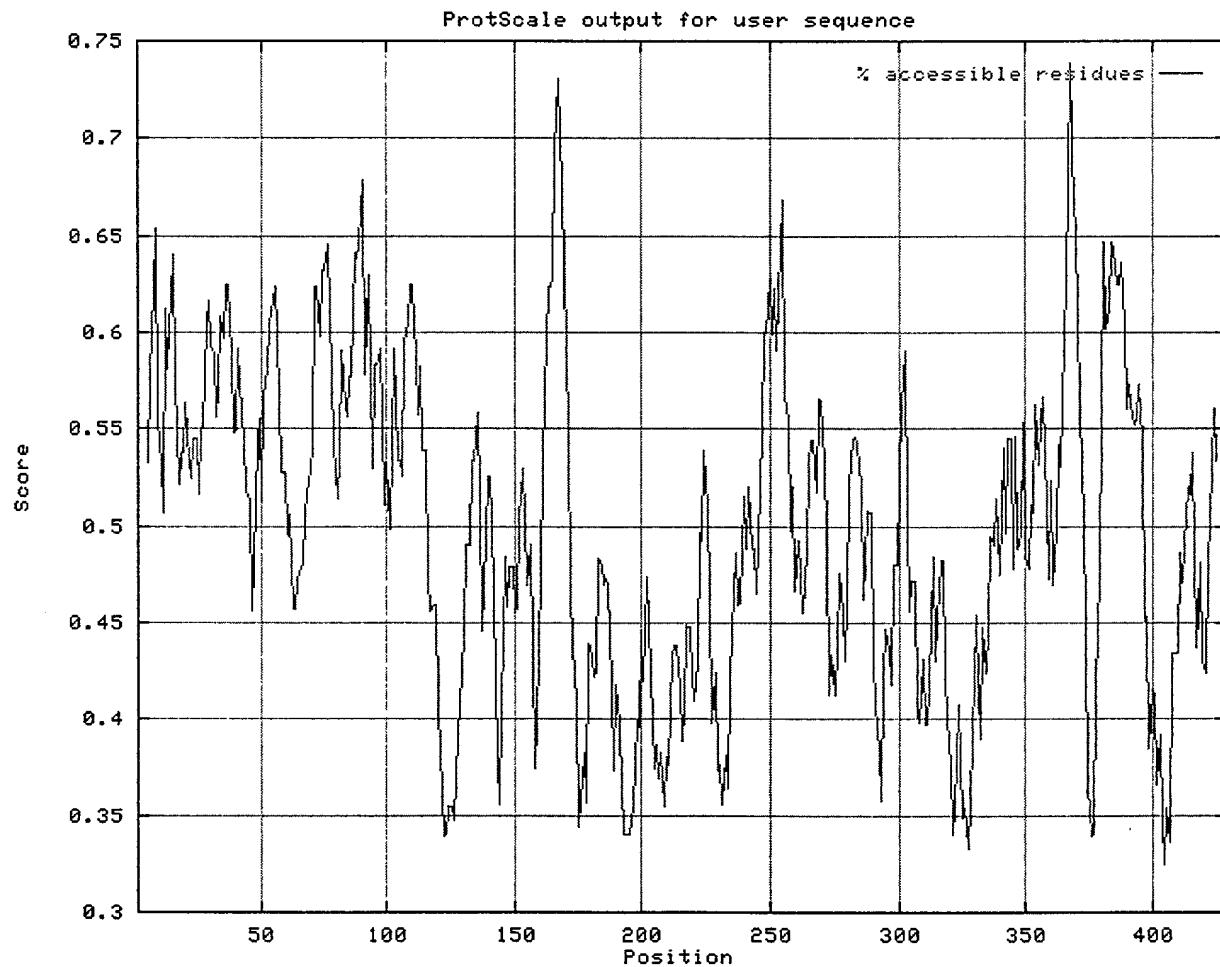


Figure 6: 108P5H8 Hydropathicity Profile  
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)



sd-70959

Figure 7: 108P5H8 % Accessible Residues Profile  
(Janin J., 1979. Nature 277:491-492)

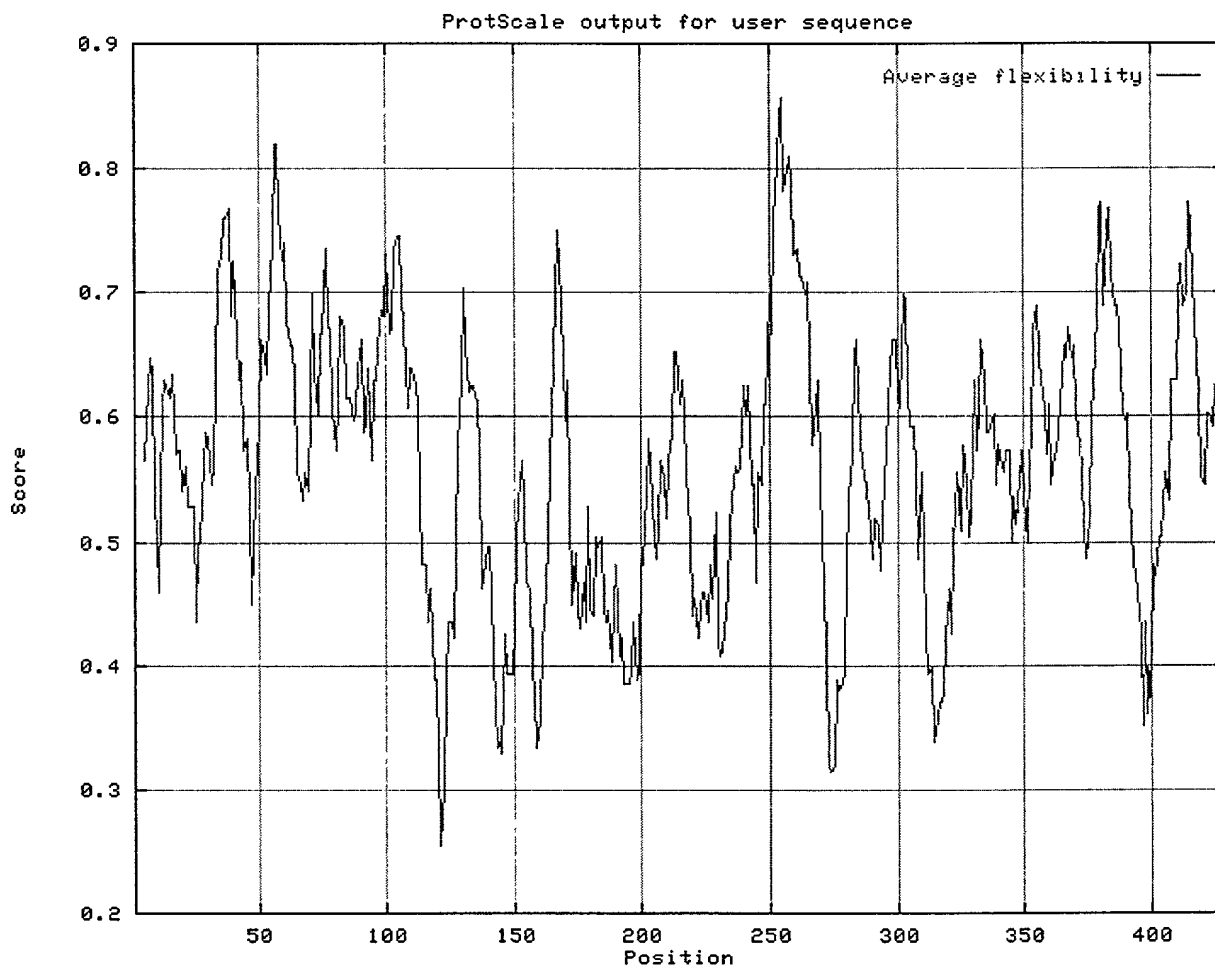


sd-70959

## Figure 8: 108P5H8 Average Flexibility Profile

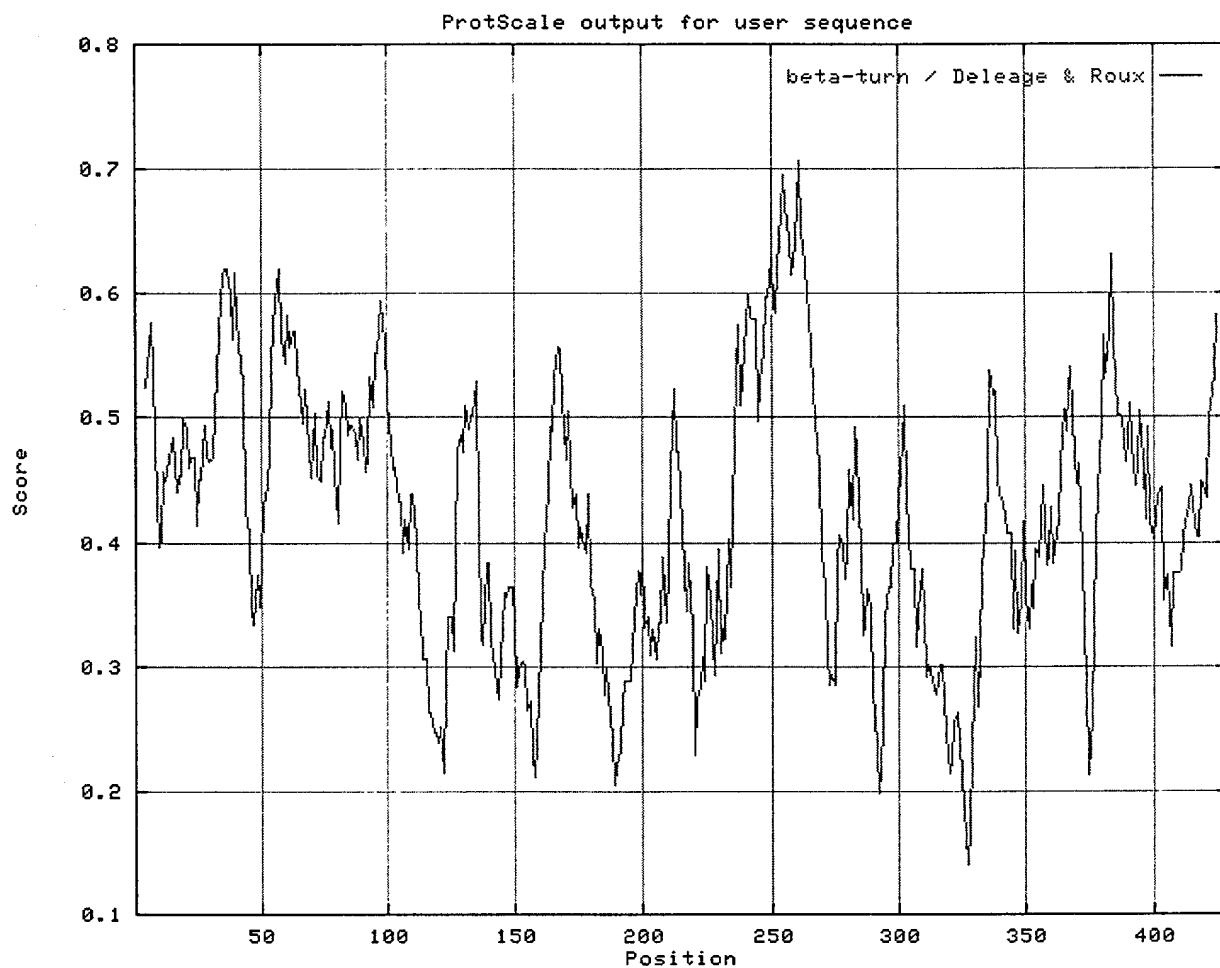
(Bhaskaran R., Ponnuswamy P.K., 1988.

Int. J. Pept. Protein Res. 32:242-255)



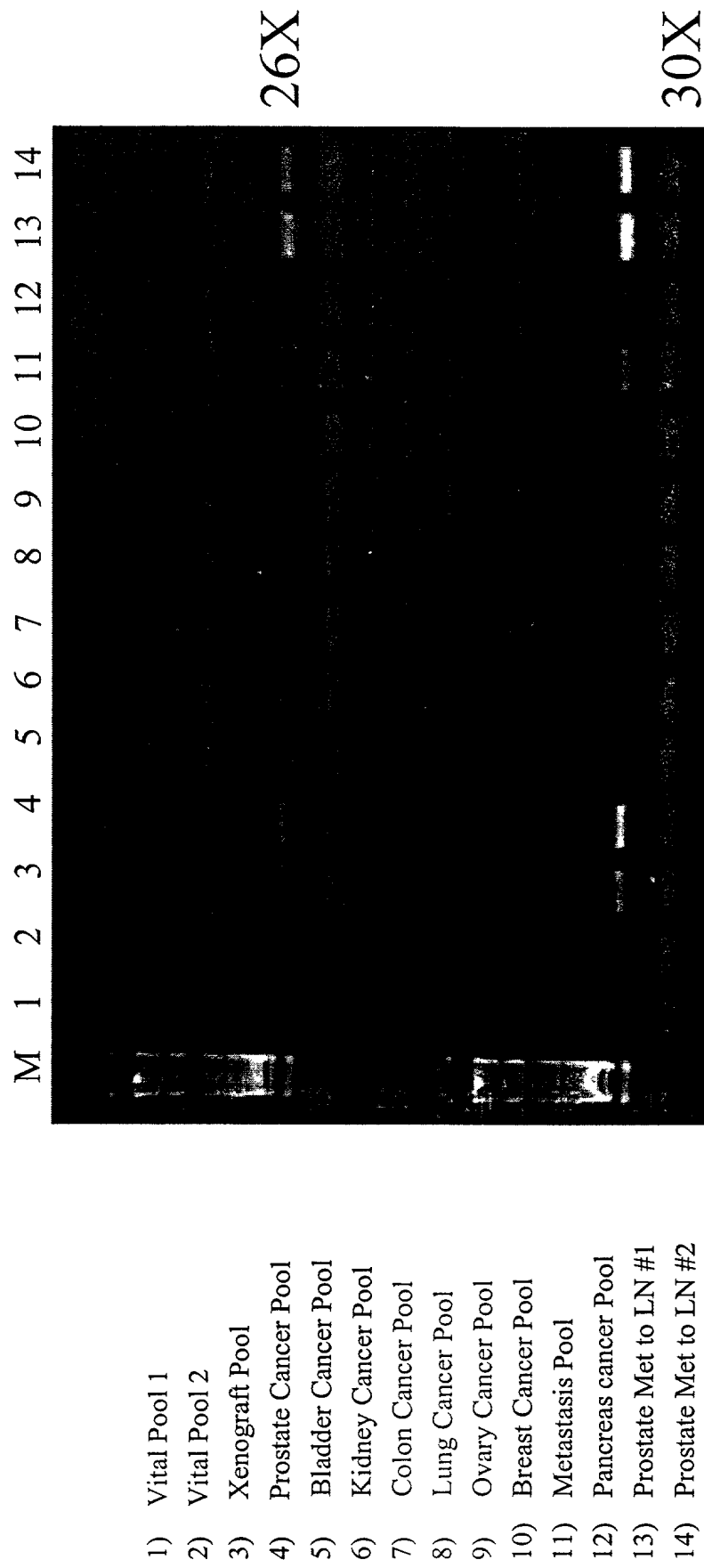
sd-70959

Figure 9: 108P5H8 Beta-turn Profile  
(Deleage, G., Roux B. 1987. Protein Engineering 1:289-294)



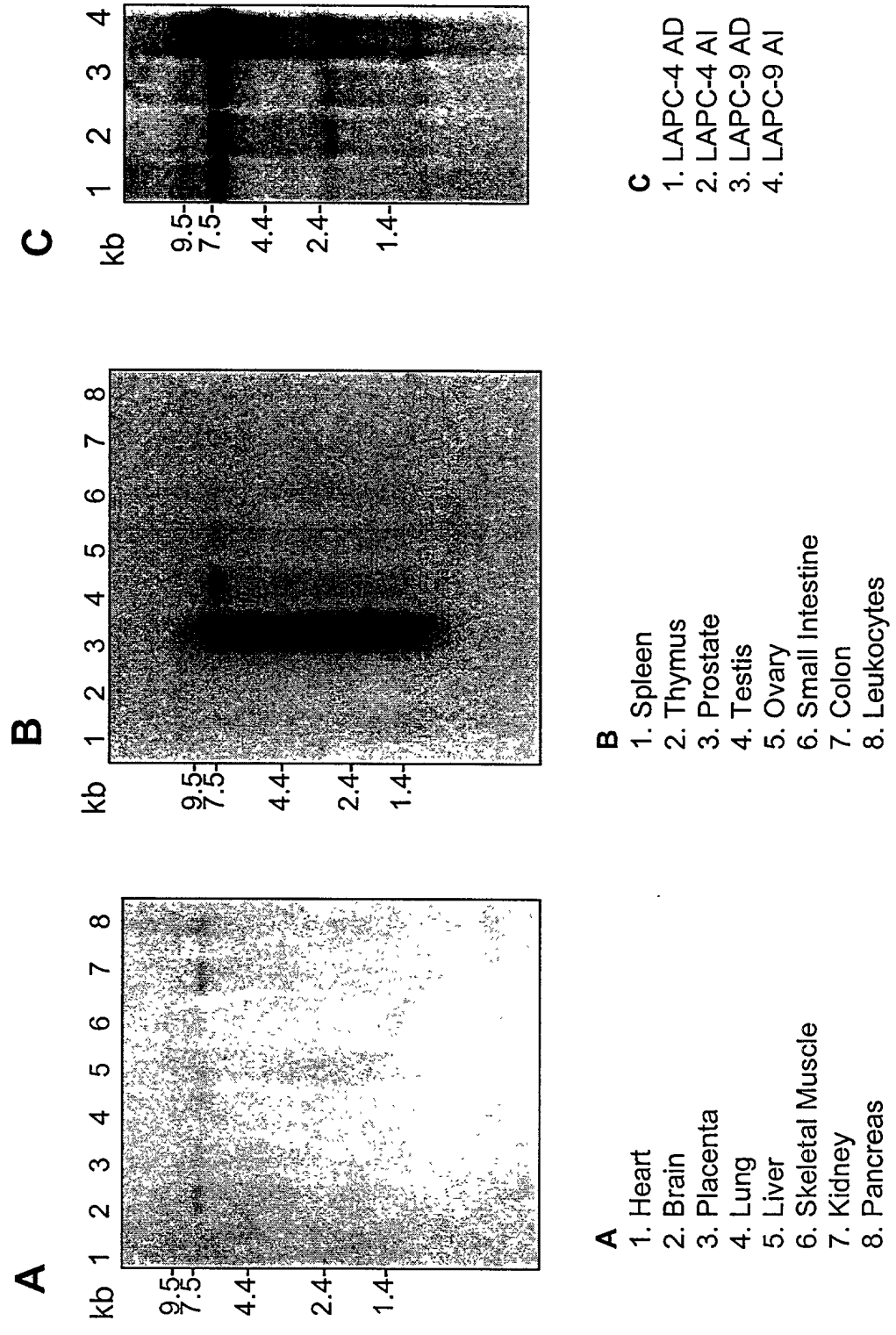
sd-70959

**Figure 10 Expression of 108P5H8 by RT-PCR**

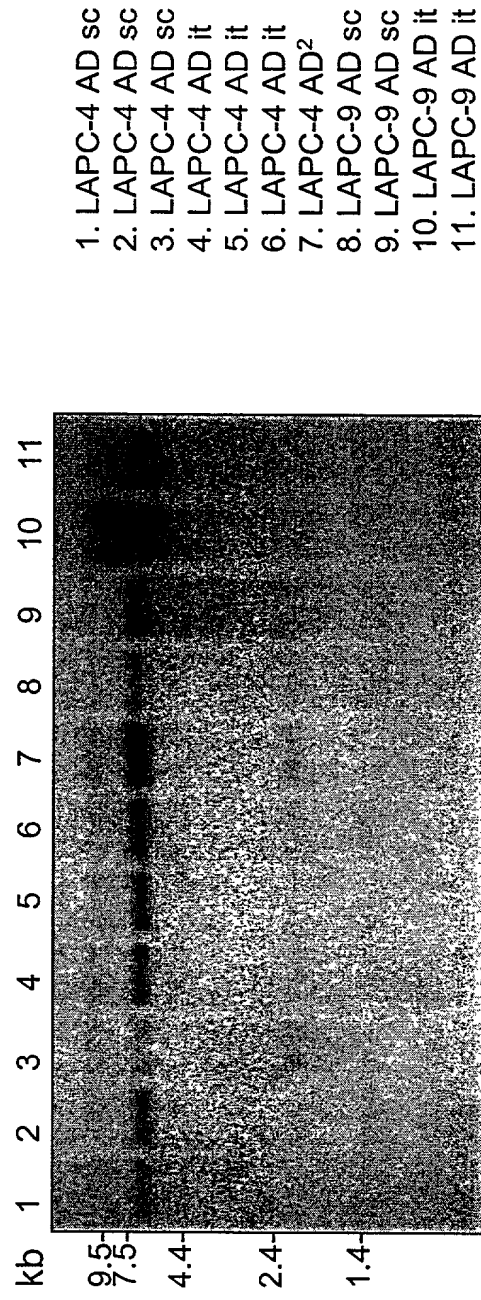




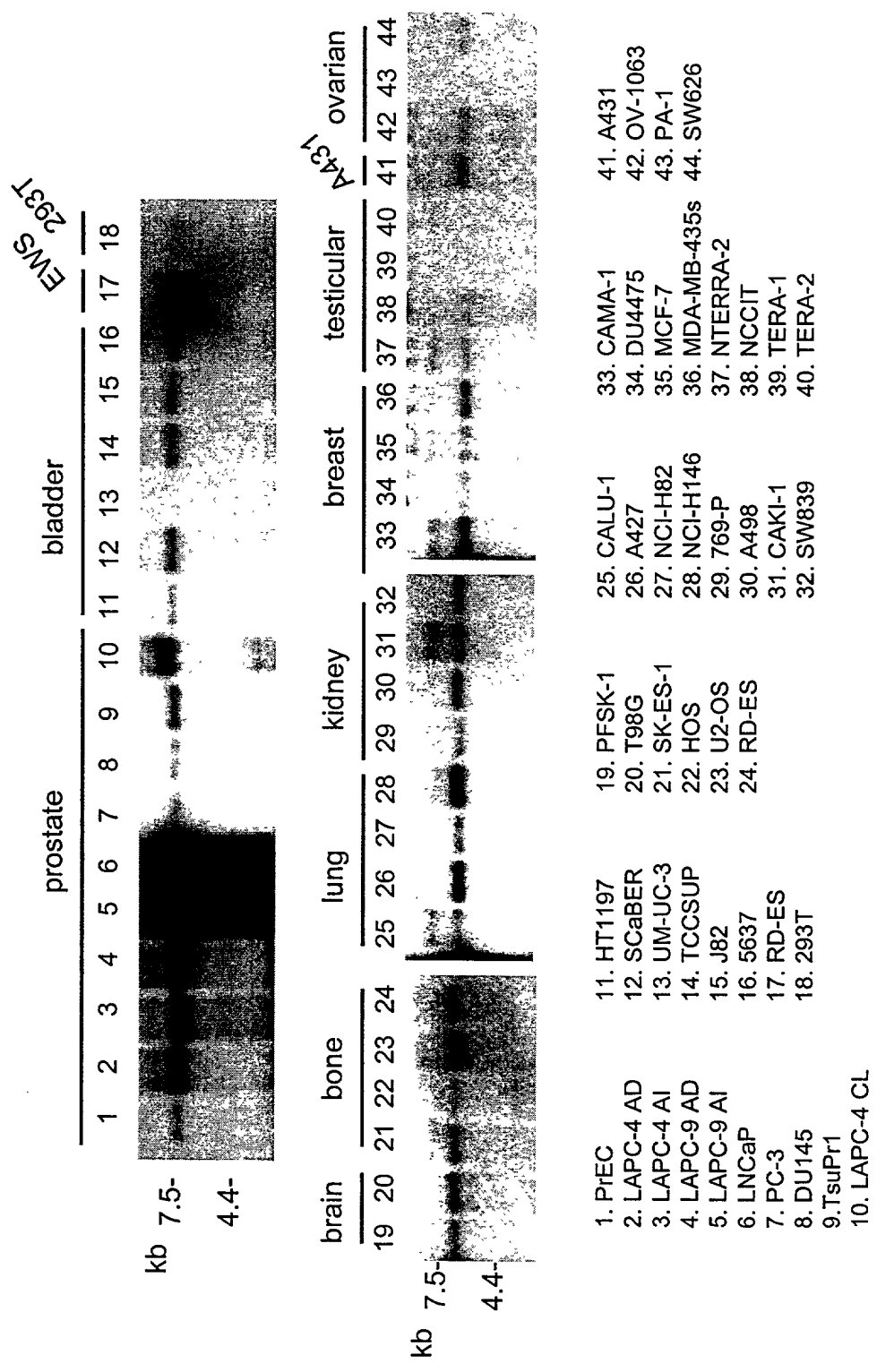
**Figure 11** Expression of 108P5H8 in Prostate Cancer Xenografts and restricted normal tissues



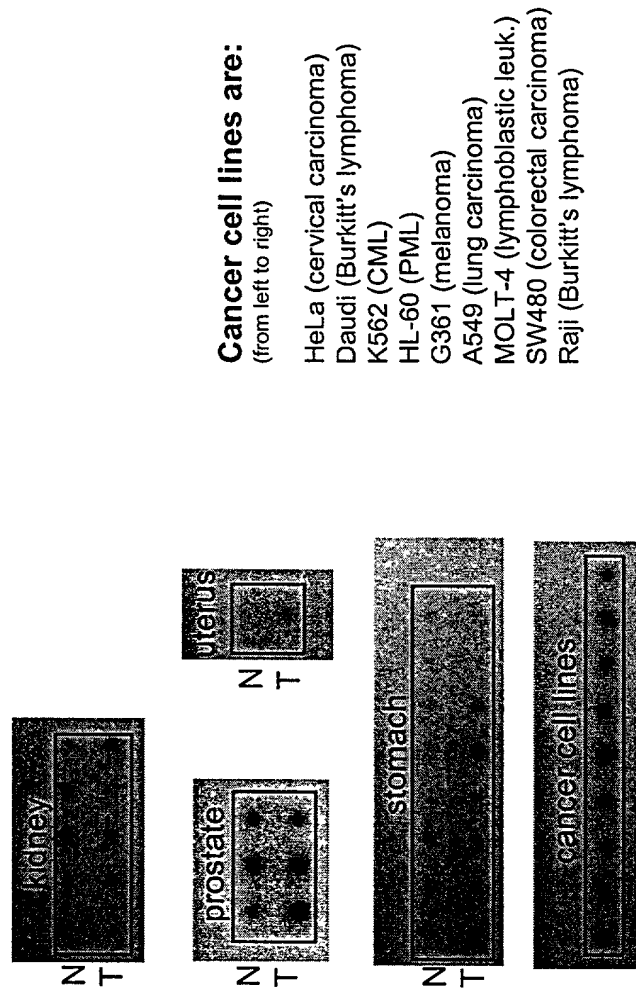
**Figure 12** Expression of 108P5H8 in prostate cancer xenografts



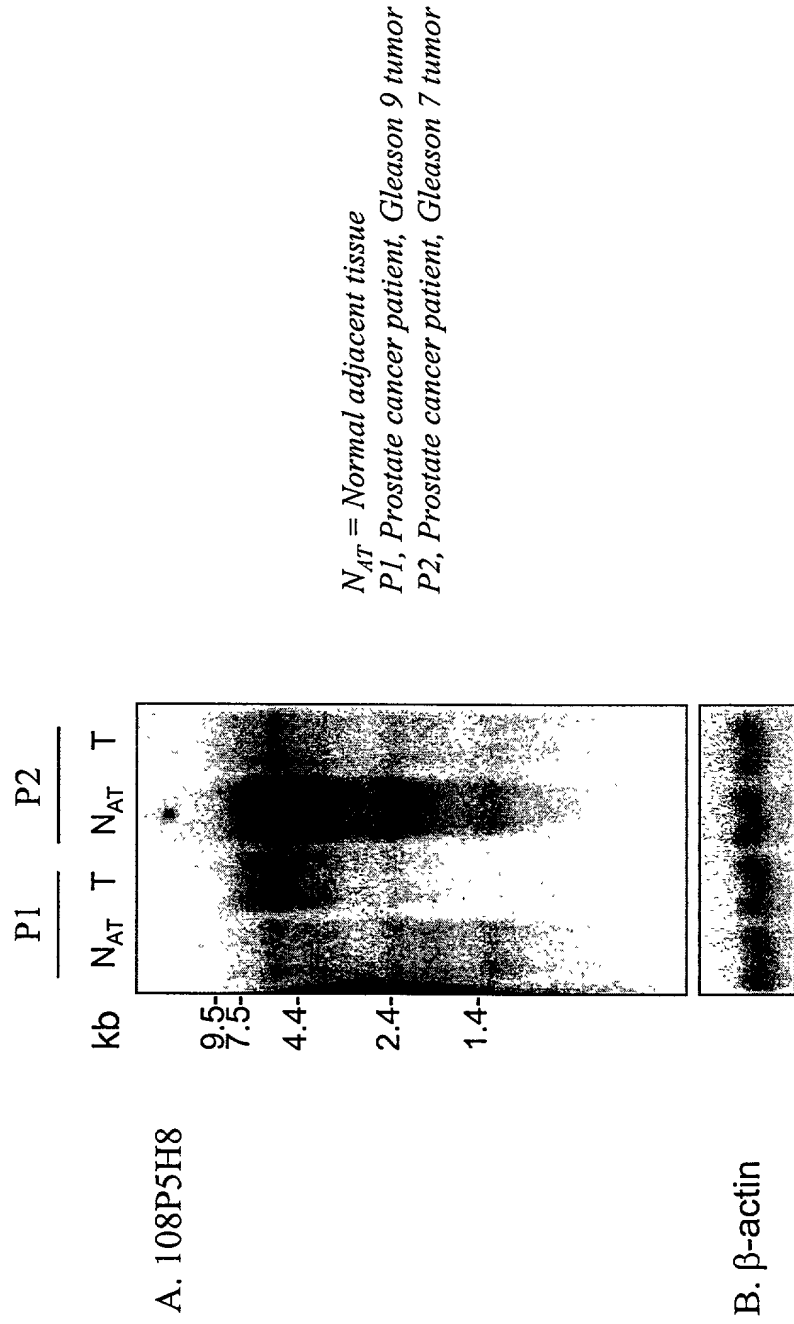
**Figure 13 Expression of 108P5H8/ZnT4 in prostate and multiple cancer cell lines**



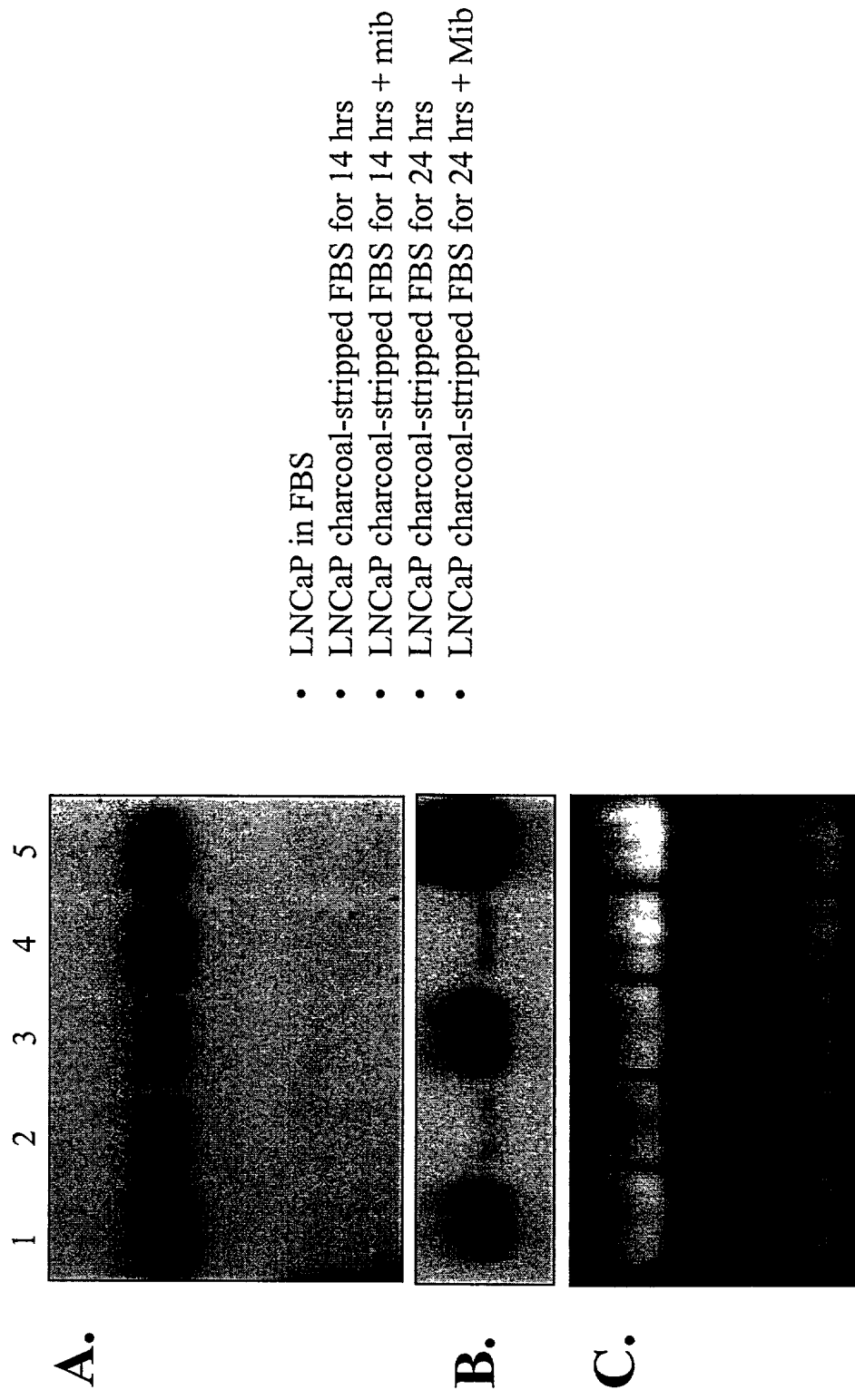
**Figure 14 Expression of 108P5H8 in Human Cancers**



**Figure 15** Expression of 108P5H8 in prostate cancer patients samples

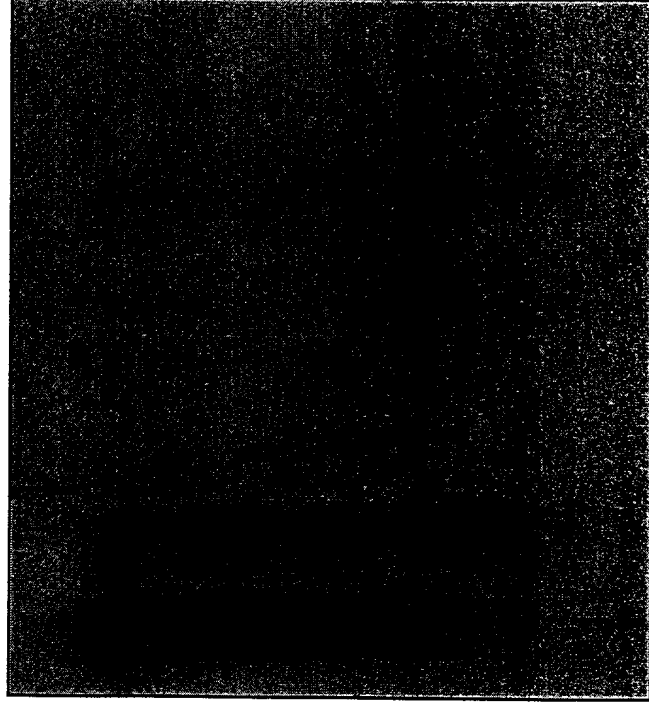


**Figure 16** 108P5H8 is not Androgen-Regulated



### **Figure 17 Expression of 108P5H8 in Prostate Cancer Metastasis Specimens**

Met1 Met2 NB NK NL NBr NO NP a



*Met1 = Prostate cancer metastasis to lymph node from patient 1*

*Met2 = Prostate cancer metastasis to lymph node from patient 2*

*NB = normal bladder*

*NK = normal kidney*

*NL = normal lung*

*NBr = normal breast*

*NO = normal ovary*

*NPa = normal pancreas*

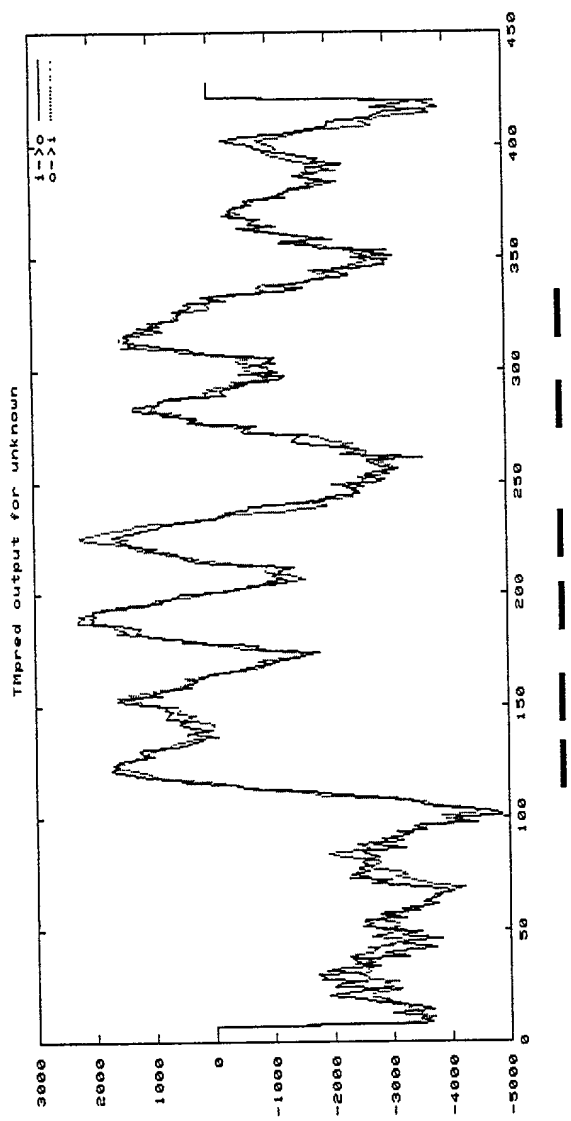
sd-70948



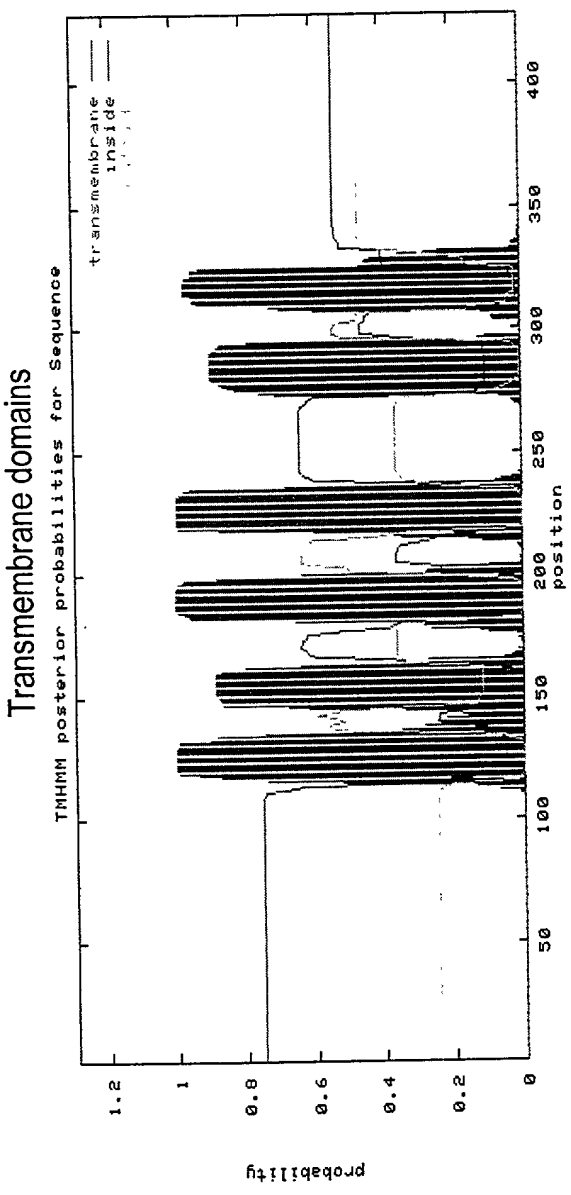
bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Figure 19: Transmembrane region prediction of 108P5H8**

**A**

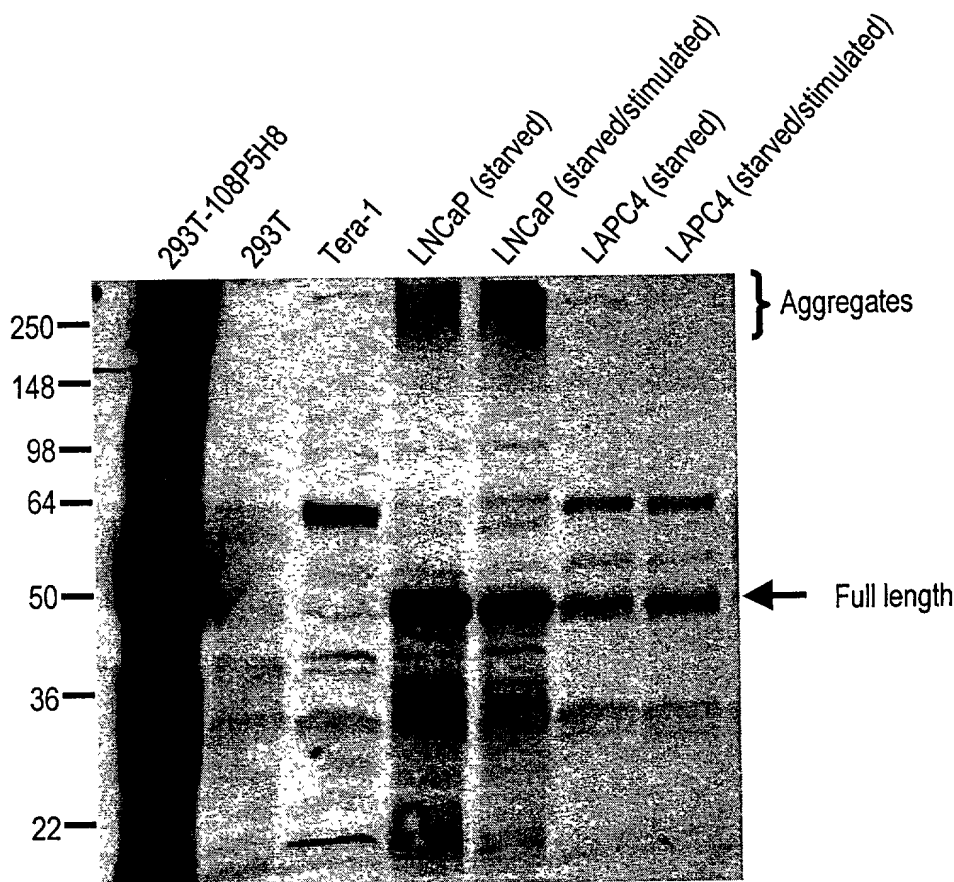


**B**

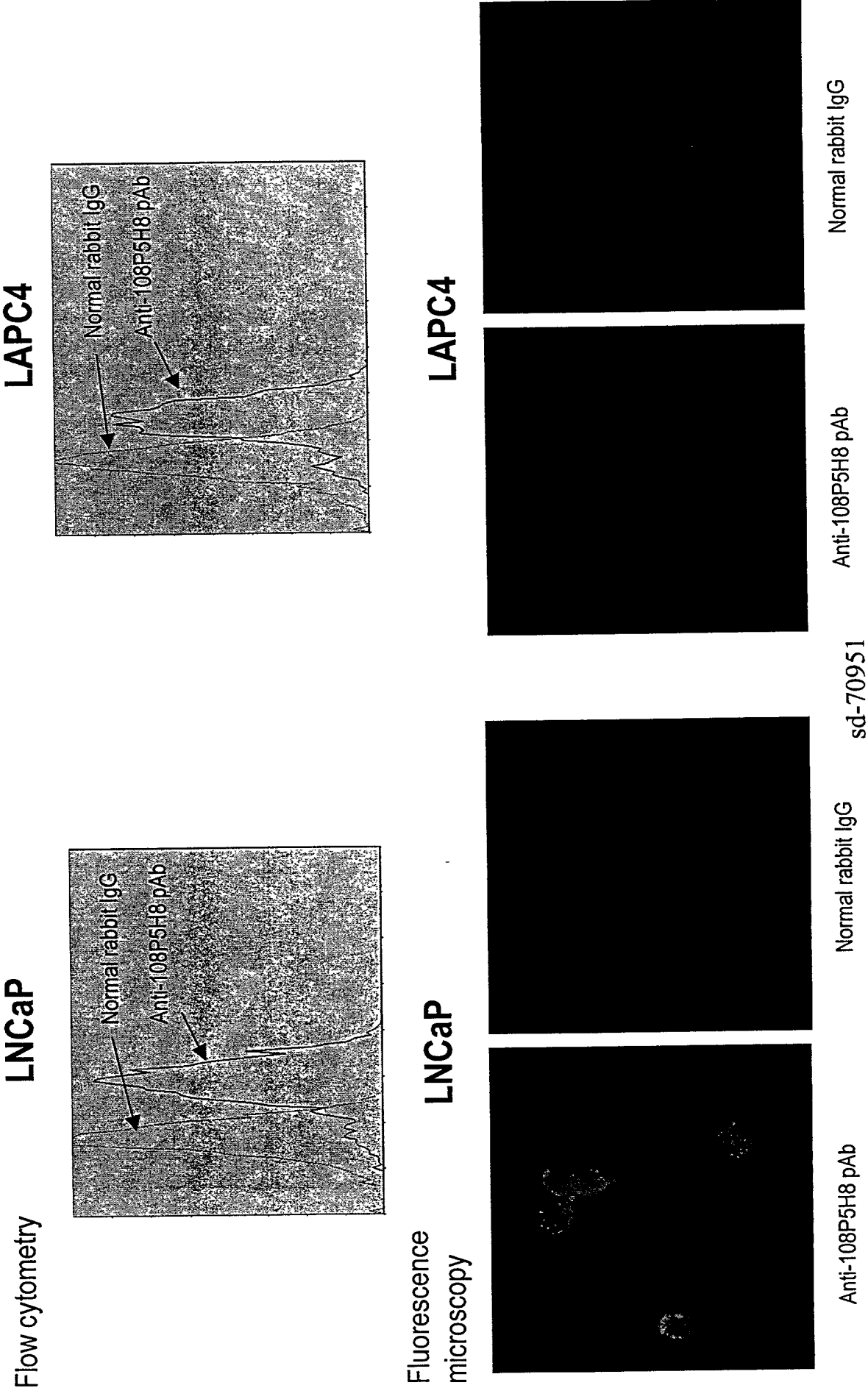


sd-70949

Figure 20: Androgen-independent expression of 108P5H8 in prostate cancer cells

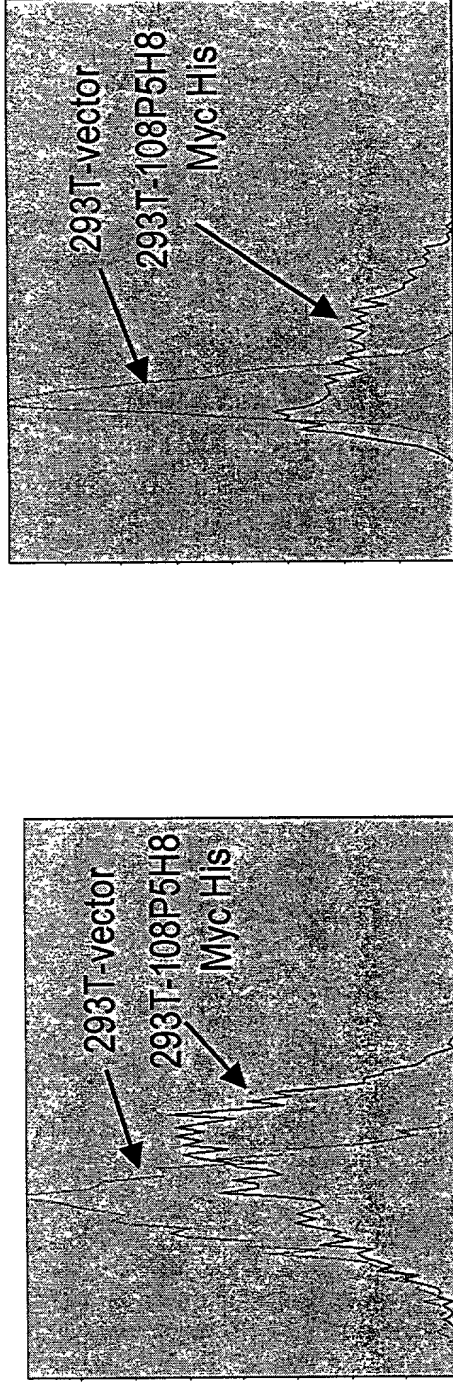


**Figure 21: Surface expression of 108P5H8 in prostate cancer cells**



## Figure 22: Surface expression of 108P5H8 in 293T cells

Flow cytometry

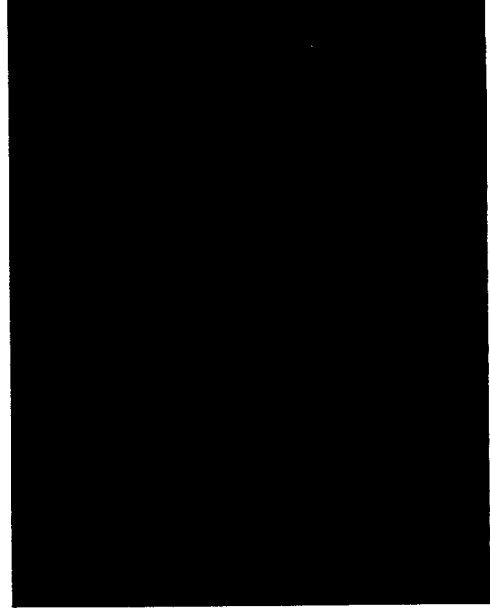


Anti-108P5H8 pAb

Fluorescence  
microscopy



293T-108P5H8

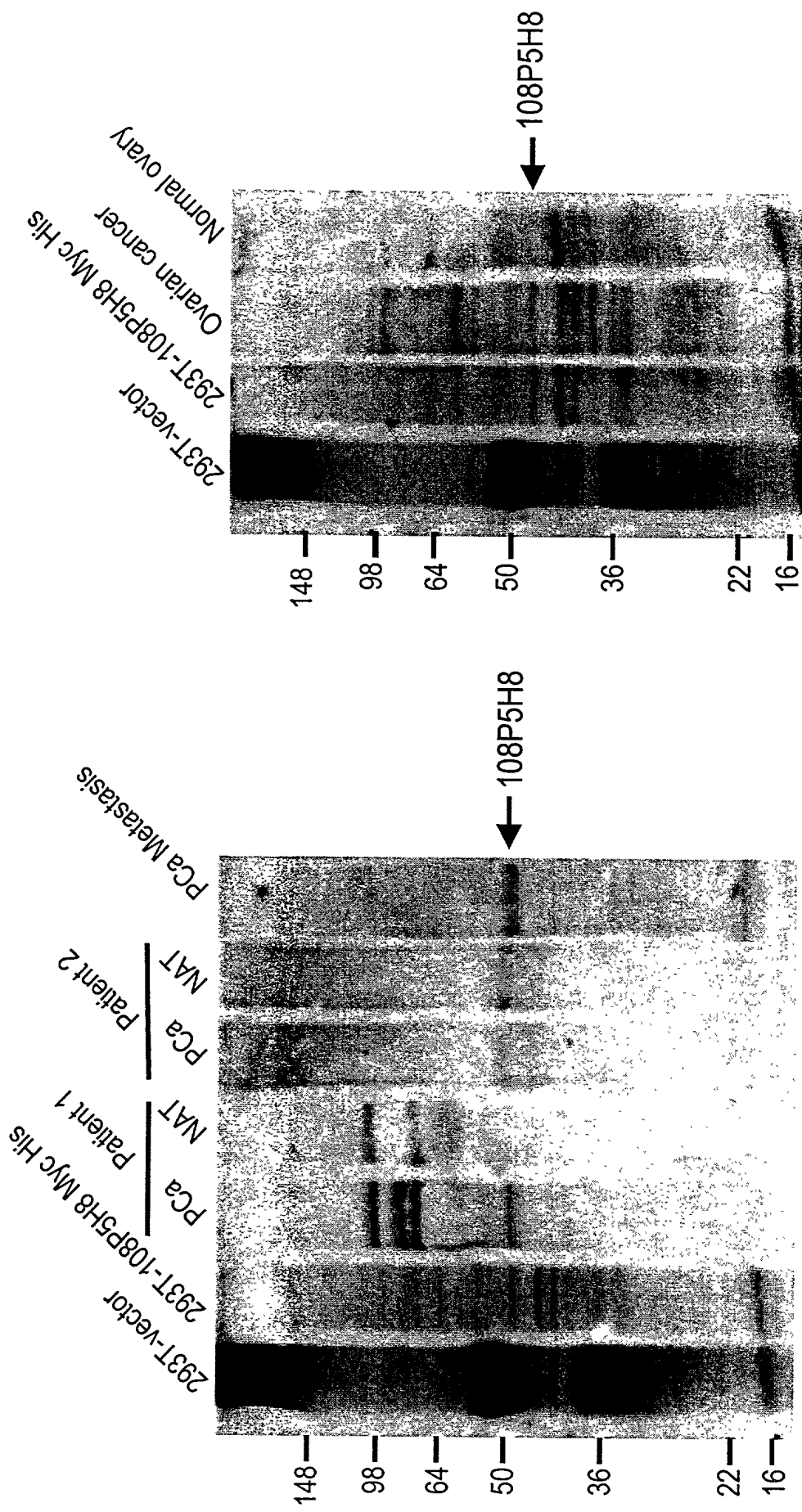


293T-vector

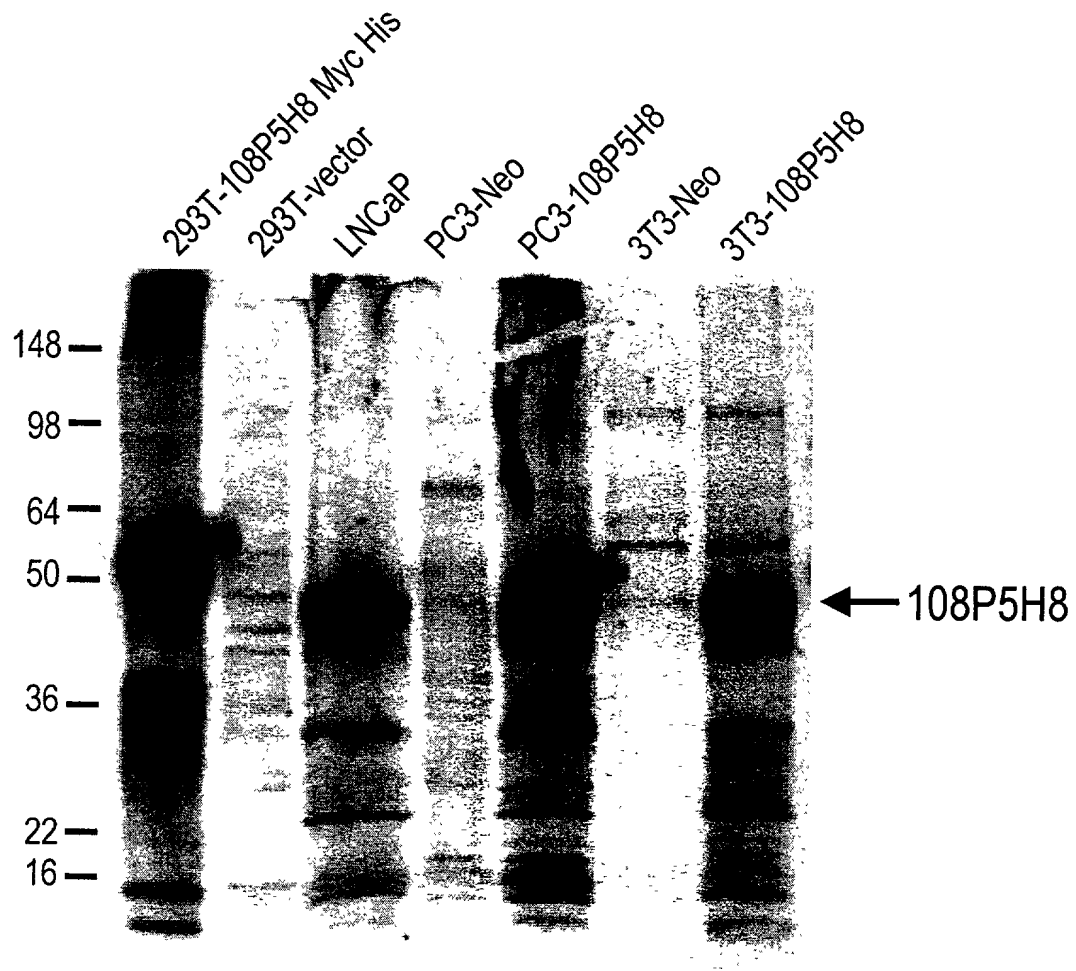
Anti-108P5H8 pAb

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332000>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 23: Expression of 108P5H8 in prostate and ovarian cancer patient specimens



**Figure 24: Expression of 108P5H8  
in engineered cell lines**



**Figure 25: Alignment of 108P5H8 v.1 protein with members of the zinc transporter family.**

A-Alignment of 108P5H8 with the human zinc transporter 4, i.e. gi 11432533.

Identities = 429/429 (100%), Positives = 429/429 (100%)

```
108P5: 1   MAGSGAWKRLKSMRLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
          MAGSGAWKRLKSMRLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER
Sbjct: 1   MAGSGAWKRLKSMRLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

108P5: 61   PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL 120
          PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL
Sbjct: 61   PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL 120

108P5: 121  YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE 180
          YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE
Sbjct: 121  YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE 180

108P5: 181  VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG 240
          VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG
Sbjct: 181  VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG 240

108P5: 241  HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
          HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE
Sbjct: 241  HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

108P5: 301  YKIADPICTYVFSLLVAFTTFRIIWDTVVILLEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
          YKIADPICTYVFSLLVAFTTFRIIWDTVVILLEGVPSHLNVDYIKEALMKIEDVYSVEDL
Sbjct: 301  YKIADPICTYVFSLLVAFTTFRIIWDTVVILLEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

108P5: 361  NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR 420
          NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR
Sbjct: 361  NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR 420

108P5: 421  CANCQSSSP 429
          CANCQSSSP
Sbjct: 421  CANCQSSSP 429
```

B-Alignment of 108P5H8 with the human zinc transporter ZNT4, i.e. gi 8134840

Identities = 428/429 (99%), Positives = 429/429 (99%)

```
108P5: 1   MAGSGAWKRLKSMRLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
          MAGSGAWKRLKSMRLRKDDAPLFLNDTSAF+FSDEAGDEGLSRFNKLRVVVADDGSEAPER
Sbjct: 1   MAGSGAWKRLKSMRLRKDDAPLFLNDTSAF+FSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

108P5: 61   PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL 120
          PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL
Sbjct: 61   PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCQREILKQRKVKAARTIAAVL 120

108P5: 121  YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE 180
          YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE
Sbjct: 121  YLLFMIGELVGGYIANS LAIMTDALHMLTDLSAI ILTLLALWLSSKSPTKRFTFGFHRLE 180

108P5: 181  VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG 240
          VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG
Sbjct: 181  VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQSG 240

108P5: 241  HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
          HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE
Sbjct: 241  HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
```

108P5: 301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIIIEGVPSHLNVDYIKEALMKIEDVYSVEDL 360  
 YKIADPICTYVFSLLVAFTTFRIIWDTVVIIIEGVPSHLNVDYIKEALMKIEDVYSVEDL  
 Sbjct: 301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIIIEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

108P5: 361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR 420  
 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR  
 Sbjct: 361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEVDR 420

108P5: 421 CANCQSSSP 429  
 CANCQSSSP  
 Sbjct: 421 CANCQSSSP 429

C-Alignment of 108P5H8 with the rat zinc transporter ZNT-4, i.e. gi 8134837

Identities = 387/430 (90%), Positives = 407/430 (94%), Gaps = 3/430 (0%)

108P5: 1 MAGSGAWKRLKSMRLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60  
 MAG GAWKRLKS+LRKDDAPLFLNDTSAFDF DE DEGLSRFNKLRVVVADD SEAPER  
 Sbjct: 1 MAGPGAWKRLKSLLRKDDAPLFLNDTSAFDFLDEVSEGLSRFNKLRVVVADDSEAPER 60

108P5: 61 PVNGAHPQLQADDDSLDDQLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120  
 PVNGAHP LQADDDSLDDQ+LPLTNSQLSLK+D CDNCSK+RE+LKQRKVK RLTTIAAVL  
 Sbjct: 61 PVNGAHPALQADDDSLDDQELPLTNSQLSLKMDPCDNCSKRRELLKQRKVKTRLTIAAVL 120

108P5: 121 YLLFMIGELVGGYIANS LAIMTDALHMLTDL SAI ILTLLALWLSSKSPTKRFTFGFHRLE 180  
 YLLFMIGELVGGY+ANSLAIMTDALHMLTDL SAI ILTLLALWLSSKSPT+RFTFGFHRLE  
 Sbjct: 121 YLLFMIGELVGGYMANSLAIMTDALHMLTDL SAI ILTLLALWLSSKSPTRRFTFGFHRLE 180

108P5: 181 VLSAMISVLLVYIILMGFLLYEAVQRTIHMNYEINGDIMLITAAGVAVNVIMGFLLNQ-- 238  
 VLSAMISV+LVY+LMGFLLYEA+QRTIHMNYEINGD+MLITAAGVAVNVIMGFLLNQ  
 Sbjct: 181 VLSAMISVMLVYVLMGFLLYEAQRTIHMNYEINGDVMLITAAGVAVNVIMGFLLNQSG 240

108P5: 239 SGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 298  
 H H+HSHSLPSNSP+ S +HGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK  
 Sbjct: 241 HHHSHAHSHSLPSNSPSMVSS-GHSHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 299

108P5: 299 PEYKIADPICTYVFSLLVAFTTFRIIWDTVVIIIEGVPSHLNVDYIKEALMKIEDVYSVE 358  
 PEYKIADPICTY+FSLLVAFTT RIIWDTVVIIIEGVPSHLNVDYIKE+LMKIEDVYSVE  
 Sbjct: 300 PEYKIADPICTYIFSLVAFTTLRIIWDTVVIIIEGVPSHLNVDYIKESLMKIEDVYSVE 359

108P5: 359 DLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNFTFGMYRCTIQLQSYRQEV 418  
 DLNIWSLTSGK+TAIVH+QLIPGSSSKWEEVQSKA HLLLNFTFGMY+CT+QLQSYRQE  
 Sbjct: 360 DLNIWSLTSGKATAIVHMLIPGSSSKWEEVQSKAKHLLLNFTFGMYKCTVQLQSYRQEAT 419

108P5: 419 RTCANCQSSS 428  
 RTCANCQSSS  
 Sbjct: 420 RTCANCQSSS 429



**Figure 26: Expression of 108P5H8 in LNCaP cells**

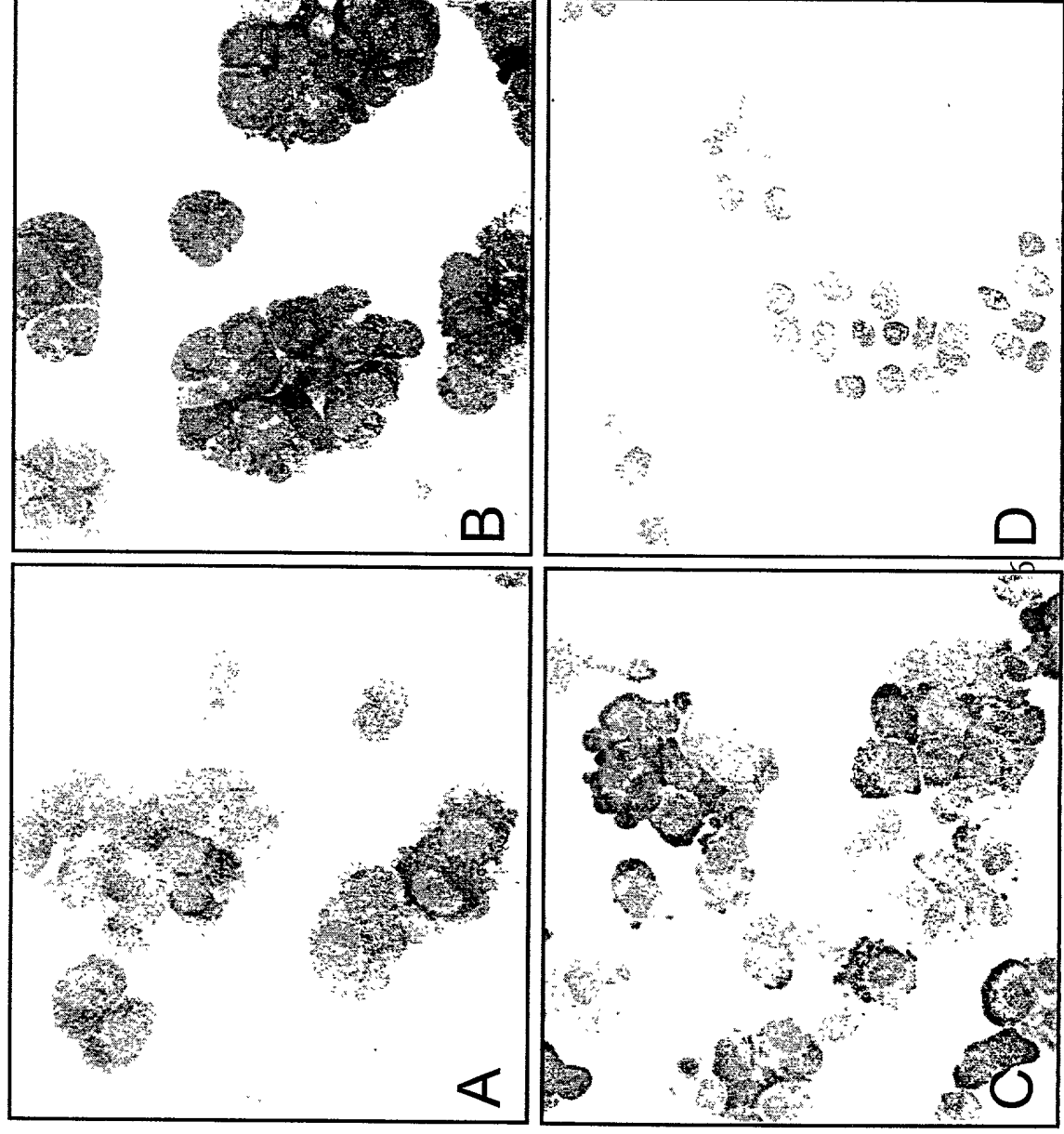
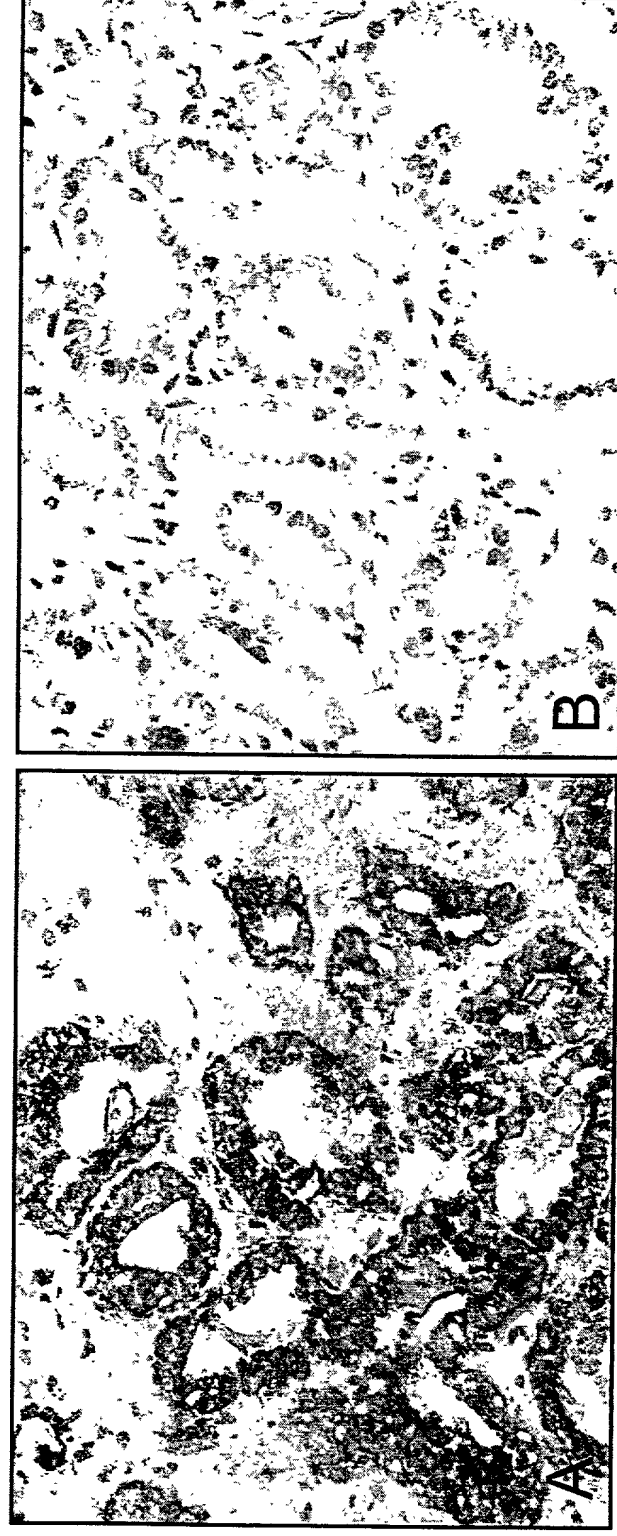


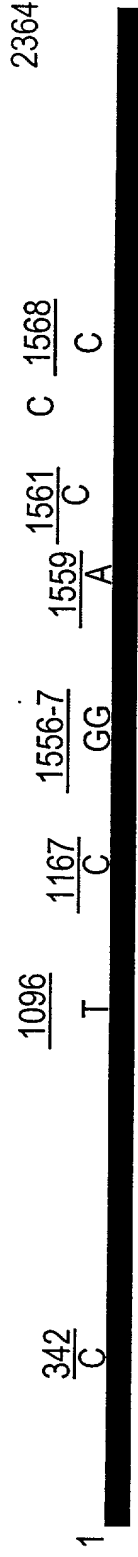
Figure 27:  
Expression of 108P5H8 in Prostate Carcinoma  
(Gleason grade 6)



**Figure 28A**

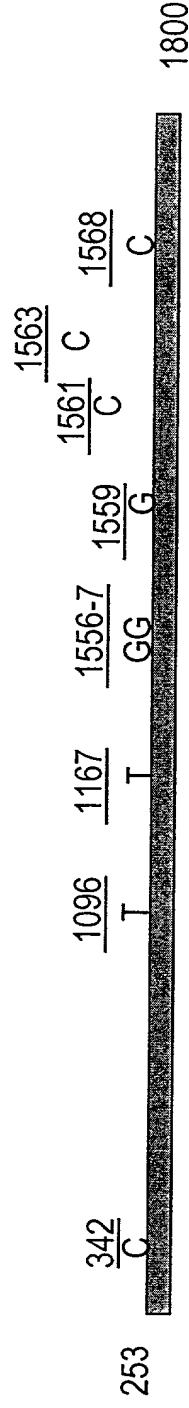
1)

108P5H08 v.1  
2364 bp



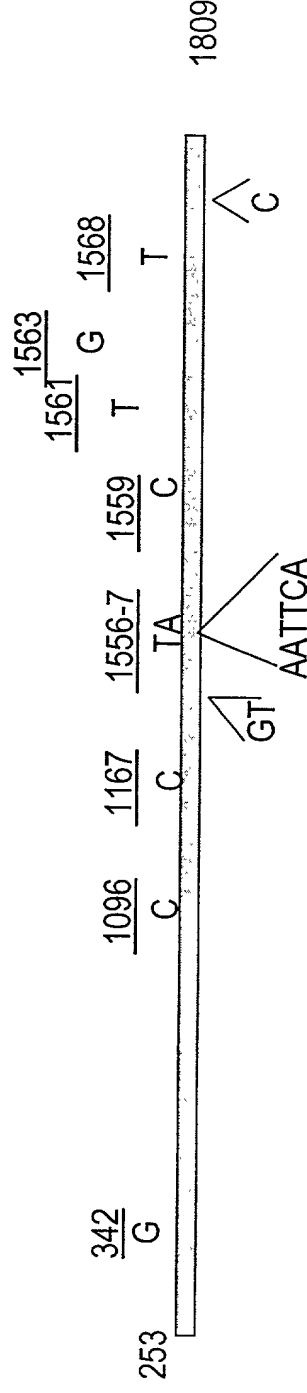
2)

108P5H08 v.2  
1548 bp



3)

108P5H08 v.3  
1557 bp\*



\*There are three insertions in variant 3: GT, AATTCA and C at 1553, 1566 and 1784, respectively; numbering is relevant to variant 1.

Figure 28B

30

D

1

1)  
108P5H08 v.1  
429 a. a.

429



30

E

1

2)  
108P5H08 v.3  
429 a.a.

429

